



**HAL**  
open science

## Sounds on time: auditory feedback in motor learning, re-learning and over-learning of timing regularity.

Floris T. van Vugt

► **To cite this version:**

Floris T. van Vugt. Sounds on time: auditory feedback in motor learning, re-learning and over-learning of timing regularity.. Neuroscience. Université Claude Bernard - Lyon I, 2013. English. NNT : . tel-00915893

**HAL Id: tel-00915893**

**<https://theses.hal.science/tel-00915893>**

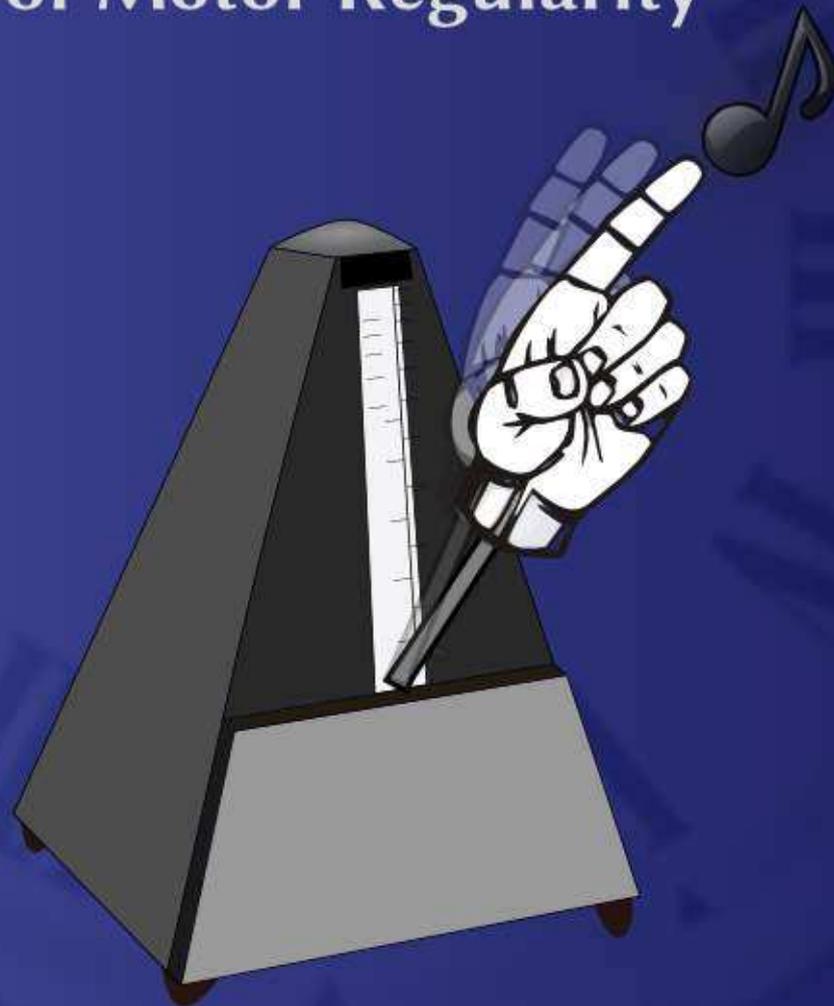
Submitted on 9 Dec 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Sounds on Time

Auditory Feedback in Learning,  
Re-Learning and Over-Learning  
of Motor Regularity



 Floris Tijmen van Vugt



N° d'ordre 232 - 2013

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE NEUROSCIENCES ET COGNITION (NSCo)

DIPLOME DE DOCTORAT EN NEUROSCIENCES

(arrêté du 7 août 2006)

Soutenue publiquement le 27 novembre 2013 à Lyon

par

Floris Tijmen VAN VUGT

Directeurs de Thèse: Dr. Barbara TILLMANN et Prof. Eckart ALTENMÜLLER

Composition du Jury

Prof. Eckart ALTENMÜLLER (directeur de thèse)

Dr. Barbara TILLMANN (directeur de thèse)

Prof. Peter KELLER (rapporteur)

Prof. Virginia PENHUNE (rapporteur)

Dr. Fabien PERRIN (président du jury)

## **Sounds on time: auditory feedback in motor learning, re-learning and over-learning of timing regularity.**

Auditory feedback is an auditory signal that contains information about performed movement. Music performance is an excellent candidate to study its influence on motor actions, since the auditory result is the explicit goal of the movement. Indeed, auditory feedback can guide online motor actions, but its influence on motor learning has been investigated less. This thesis investigates the influence of auditory feedback in motor learning, focusing particularly on how we learn temporal control over movements. First, we investigate motor learning in non-musicians, finding that they benefit from temporal information supplied by the auditory signal and are sensitive to distortions of this temporal information. Second, we turn to stroke patients that are re-learning motor actions in a rehabilitation setting. Patients improved their movement capacities but did not depend on the time-locking between movements and the resulting auditory feedback. Surprisingly, they appear to benefit from distortions in feedback. Third, we investigate musical experts, who arguably have established strong links between movement and auditory feedback. We develop a novel analysis framework that allows us to segment timing into systematic and non-systematic variability. Our finding is that these experts have become largely independent of the auditory feedback. The main claim defended in this thesis is that auditory feedback can and does play a role in motor learning of regularity, but the way in which it is used varies qualitatively between different populations. These findings provide new insights into auditory-motor integration and are relevant for developing new perspectives on the role of music in training and rehabilitation settings.

### **Keywords:**

Sensorimotor Integration, Auditory Feedback, Motor Learning, Motor regularity, Timing, Rehabilitation, Expert musicians, Sequence Production, Sequence Learning

## **Feedback auditif et régularité motrice : apprentissage, réhabilitation et expertise**

Le feedback auditif se définit comme un signal auditif qui contient de l'information sur un mouvement. Il a été montré que le feedback auditif peut guider le mouvement en temps réel, mais son influence sur l'apprentissage moteur est moins clair. Cette thèse a pour but d'examiner l'influence du feedback auditif sur l'apprentissage moteur, en se focalisant sur le contrôle temporel des mouvements. Premièrement, nous étudions l'apprentissage moteur chez les non-musiciens sains et montrons qu'ils bénéficient de l'information temporelle contenue dans le feedback auditif et qu'ils sont sensibles aux distortions de cette information temporelle. Deuxièmement, nous appliquons ces connaissances à la réhabilitation de patients cérébro-lésés. Nous trouvons que ces patients améliorent leurs capacités de mouvement mais ne dépendent pas de la correspondance temporelle entre le mouvement et le son. Paradoxalement, ces patients ont même bénéficié des distortions temporelles dans le feedback. Troisièmement, nous étudions les experts musicaux, car ils ont établi des liens particulièrement forts entre leur mouvement et le son. Nous développons de nouveaux outils d'analyse qui nous permettent de séparer les déviations temporelles en variation systématique et non-systématique. Le résultat principal est que ces experts sont devenu largement indépendents du feedback auditif. La proposition centrale de cette thèse est que le feedback auditif joue un rôle dans l'apprentissage moteur de la régularité, mais la façon dont le cerveau l'utilise dépend de la population étudiée. Ces résultats donnent une nouvelle perspective sur l'intégration audio-motrice et contribuent au développement de nouvelles approches pour l'apprentissage de la musique et la réhabilitation.

### **Mots clés:**

Intégration sensorimotrice, feedback auditif, apprentissage moteur, régularité motrice, timing, réhabilitation, experts musicaux, production de séquences, apprentissage de séquences.

**Institutes:**

Institute of Music Physiology and Musicians' Medicine,  
University of Music, Drama and Media,  
Emmichplatz 1, 30175 Hanover  
Germany

Lyon Neuroscience Research Center  
CNRS-UMR 5292, INSERM U1028,  
University Lyon-1,  
50 av Tony Garnier, 69007 Lyon,  
France

This research was supported by the EBRAMUS  
(European Brain and Music) grant, ITN MC FP7, GA 238157.

The musician may sing to you of the rhythm which is in all space,  
but he cannot give you the ear which arrests the rhythm nor the  
voice that echoes it.

– Kahlil Gibran, *The Prophet*

Sounds on time -

## Table of Contents

<b>I. Sensorimotor integration: a conceptual map.....</b>	<b>19</b>
<b>I.1. Introduction.....</b>	<b>20</b>
1.1 Aim.....	20
1.2 Disclaimer: finding a balance between unity and diversity.....	21
<b>I.2. Preliminaries.....</b>	<b>22</b>
2.1 The human motor system – a brief overview.....	23
2.1.a Motor control areas – a brief outline.....	23
2.1.b Control Strategies.....	26
<b>I.3. Towards a definition of sensorimotor integration.....</b>	<b>28</b>
3.1 Sensorimotor integration.....	28
3.2 Aim.....	29
<b>I.4. Motor theories of cognition.....</b>	<b>31</b>
<b>I.5. Motor involvement in perception.....</b>	<b>32</b>
5.1 Movement making perception possible.....	34
5.1.a Microsaccades (vision).....	35
5.1.b Motor-induced suppression.....	36
5.2 Action perception.....	39
5.2.a Mirror neurons.....	39
5.2.b Mirror networks.....	43
5.2.c Direct matching hypothesis and motor resonance.....	46
5.2.d Plasticity of mirror networks: arbitrariness of the associations.....	47
5.3 Speech perception and comprehension.....	50
5.3.a Perception of speech sounds.....	50
5.3.b Syntactic processing.....	54

Sounds on time -

5.3.c Semantic processing.....	55
5.3.d Conversational interaction.....	56
5.3.e Anatomical substrates.....	57
5.3.f Necessity and role of motor involvement.....	57
<b>5.4 Music perception.....</b>	<b>58</b>
5.4.a Why music?.....	58
5.4.b Motor systems supporting perceptual memory formation.....	59
5.4.c Music perception in novices: a focus on rhythm perception.....	61
5.4.d Musical experts.....	62
5.4.e Creation of auditory-motor associations: music learning in novices.....	64
<b>I.6. Sensory involvement in action.....</b>	<b>66</b>
<b>6.1 Sensory feedback in action.....</b>	<b>70</b>
6.1.a Representing action by its sensory consequences (action selection).....	70
6.1.b Sensory guidance of action (action execution).....	72
6.1.c Motor learning through sensory feedback (action learning).....	73
<b>6.2 Sensory stimulation in action.....</b>	<b>75</b>
6.2.a Sensorimotor corticospinal facilitation.....	75
6.2.b Sensation as a reference for movement.....	77
<b>6.3 Speech production.....</b>	<b>79</b>
6.3.a Sensory co-activation.....	79
6.3.b Reliance on feedback.....	79
<b>6.4 Music production.....</b>	<b>80</b>
6.4.a Sensory co-activation during silent performance.....	81
6.4.b Reliance on feedback.....	82
6.4.c Sonification feedback in musical novices.....	82
<b>I.7. Synthesis.....</b>	<b>84</b>
<b>I.8. Outline of the thesis.....</b>	<b>85</b>
<b>II. Learning in musical novices.....</b>	<b>87</b>
<b>II.1. Introduction.....</b>	<b>88</b>
1.1 Motor regularity.....	88

1.2 Serial Reaction Time.....89  
1.3 Aims and Hypotheses.....91

**II.2. Thresholds of auditory-motor coupling measured with a simple task in musicians and non-musicians: Was the sound simultaneous to the keystroke?. 92**

2.1 Introduction.....94  
2.2 Materials and Methods.....96  
    2.2.a Participants.....96  
    2.2.b Materials.....99  
    2.2.c Procedure.....100  
2.3 Results.....103  
    2.3.a Delay Detection.....103  
    2.3.b Anisochrony.....104  
    2.3.c Synchronisation-Continuation Tapping.....105  
    2.3.d Comparison between the tests.....106  
2.4 Conclusion.....108

**II.3. Auditory feedback benefits short term motor learning.....116**

3.1 Introduction.....119  
3.2 Methods.....125  
    3.2.a Participants.....125  
    3.2.b Materials.....125  
    3.2.c Procedure.....128  
    3.2.d Data Analysis.....134  
3.3 Results.....137  
    3.3.a Sequence Learning.....137  
    3.3.b Sequence switching.....138  
    3.3.c Muting effects.....140  
    3.3.d Scale tapping.....141  
    3.3.e Anisochrony detection task.....142  
    3.3.f Delay detection task.....142  
    3.3.g Synchronisation-Continuation Tapping.....142  
3.4 Discussion.....144  
    3.4.a Sequence learning and auditory feedback.....144

Sounds on time -

3.4.b Sequence-specific vs. unspecific learning.....	147
3.4.c Auditory feedback deprivation.....	148
3.4.d Scale playing transfer task.....	149
3.4.e Does the sound need to be time-locked to the movement?.....	150
3.4.f Anisochrony, delay detection and synchronisation-continuation tapping tasks.....	151
3.4.g Limitations and outlook.....	152
3.5 References.....	152
3.6 Acknowledgments.....	155
<b>II.4. Discussion.....</b>	<b>162</b>
<b>III. Re-learning in stroke patients.....</b>	<b>163</b>
<b>III.1. Introduction.....</b>	<b>164</b>
1.1 Stroke: pathophysiology and rehabilitation.....	164
1.1.a Causes of motor impairments.....	164
1.1.b Mechanisms underlying neural repair after stroke.....	166
1.2 Proprioception and motor (re)learning.....	167
1.2.a What is proprioception?.....	168
1.2.b Proprioception tests.....	168
1.2.c Alterations in proprioception.....	170
1.2.d Proprioception and motor learning.....	170
1.2.e Proprioception after stroke and its role in rehabilitation.....	171
1.3 Music-Supported Therapy: working hypotheses.....	172
1.3.a Music interventions in clinical settings.....	172
1.3.b Music-supported therapy.....	172
1.3.c Probing the mechanisms of music-supported therapy.....	173
1.4 Aims and hypotheses.....	175
<b>III.2. Random delay is more efficient than immediate auditory feedback in fine motor rehabilitation after stroke.....</b>	<b>178</b>
2.1 Introduction.....	181
2.2 Methods.....	184
2.2.a Patient population.....	184

Sounds on time -

2.2.b Therapy.....	187
2.2.c Nine-hole pegboard test.....	188
2.2.d Finger tapping measurements.....	188
2.2.e Patient mood measurements.....	190
2.2.f Auditory and sensorimotor tasks.....	191
2.2.g Intra-therapy measurements.....	193
2.2.h Qualitative interview.....	194
2.2.i Statistical analyses.....	194
2.2.j Ethics.....	194
<b>2.3 Results.....</b>	<b>195</b>
2.3.a Nine hole pegboard test.....	195
2.3.b Finger tapping measurements.....	195
2.3.c Patient mood measurements (profile of mood states).....	201
2.3.d Faces scale.....	201
2.3.e Auditory and auditory-motor tests.....	202
2.3.f Within-therapy measure: keyboard keystrokes.....	203
2.3.g Qualitative interviews.....	204
<b>2.4 Discussion.....</b>	<b>205</b>
<b>2.5 References.....</b>	<b>212</b>
<b>2.6 Conflict of interest statement.....</b>	<b>215</b>
<b>2.7 Acknowledgements.....</b>	<b>215</b>

**III.3. Music-supported motor training after stroke reveals no superiority of synchronisation in group therapy.....220**

<b>3.1 Introduction.....</b>	<b>222</b>
<b>3.2 Methods.....</b>	<b>223</b>
3.2.a Patient group characteristics.....	224
3.2.b Music training.....	225
3.2.c Nine-hole pegboard test.....	226
3.2.d Finger tapping measurements.....	227
3.2.e Mood test: Profile of Mood States.....	229
3.2.f Mood test: faces scale.....	229
3.2.g Ethics.....	229
3.2.h Data analysis.....	230

3.3 Results.....	230
3.3.a Nine-hole pegboard test.....	230
3.3.b Finger tapping tests.....	231
3.3.c Mood tests.....	235
3.3.d Faces scale mood ratings.....	236
3.4 Conclusions.....	238
<b>III.4. Discussion.....</b>	<b>244</b>
4.1 Auditory feedback.....	244
4.2 A note about mood improvements.....	245
4.3 Implications for clinical practice.....	245
<b>IV. Over-learning in Musical Experts.....</b>	<b>247</b>
<b>IV.1. Introduction.....</b>	<b>248</b>
1.1 Choice of experimental paradigm.....	248
1.2 Piano timing measurements: methodological considerations.....	249
1.3 Musical performance and the perception-action interface.....	251
1.4 Musical scales.....	252
1.4.a Description of the scale playing movement.....	252
1.4.b Timing control in scale playing.....	253
1.4.c Towards a new model of timing in piano performance.....	256
1.5 Aims and hypotheses.....	258
<b>IV.2. Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking.....</b>	<b>260</b>
<b>IV.3. Individuality that is unheard of: systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint.....</b>	<b>270</b>
<b>IV.4. Spatial and temporal symmetries of motor primitives in skilled piano performance at different tempi.....</b>	<b>282</b>
4.1 Introduction.....	284
4.2 Methods.....	287

Sounds on time -

4.2.a	Participants.....	287
4.2.b	Procedure.....	287
4.2.c	Processing of MIDI data.....	288
4.2.d	Scale unevenness.....	288
4.2.e	Unevenness across tempi.....	289
4.2.f	Individual note timing.....	289
4.2.g	Towards a model of timing deviations.....	289
4.2.h	Fitting methods.....	290
4.3	Results.....	291
4.3.a	Scale-level tempo dependencies.....	291
4.3.b	Note-by-note irregularities: establishing the phrasal and neuromuscular templates	293
4.3.c	Trade-off between the two templates.....	295
4.3.d	Generalising to all outward scales.....	297
4.3.e	Generalising to inward scales.....	297
4.3.f	Generalising to both hands and directions.....	299
4.3.g	Individual level and correlation with accumulated playing time.....	299
4.4	Discussion.....	301
4.5	References.....	307
<b>IV.5</b>	<b>Discussion.....</b>	<b>318</b>
5.1	Outline of the irregularity-instability model.....	318
5.2	Auditory feedback.....	320
5.3	Future studies.....	322
<b>V</b>	<b>General Discussion.....</b>	<b>325</b>
<b>V.1</b>	<b>Auditory feedback is implicated in motor learning, but how depends on the population.....</b>	<b>326</b>
1.1	Disruptions in auditory-motor coupling: when it helps and when it hurts.....	327
1.1.a	Healthy non-musicians: sound feedback-based learning.....	327
1.1.b	Stroke patients: learning not to rely on sound feedback.....	329
1.1.c	Musical experts: learning through feedback, performance through feedforward.....	337
<b>V.2</b>	<b>Sensorimotor integration: toward a unified approach.....</b>	<b>337</b>

2.1 From a heterogeneous account.....	337
2.2 ... to a unified account of sensorimotor integration.....	338
<b>VI. References and indices.....</b>	<b>343</b>
<b>VII. Appendix: Parkinson's Disease patients.....</b>	<b>371</b>
<b>VII.1. Effects of dopaminergic and subthalamic stimulation on musical performance.....</b>	<b>372</b>
<b>VIII. Appendix: Fine-tuning the timing analysis in experts.....</b>	<b>379</b>
<b>VIII.1. The influence of chronotype on making music: Circadian fluctuations in pianists' fine motor skills.....</b>	<b>380</b>
<b>VIII.2. Musician's dystonia in pianists: long-term evaluation of retraining and other therapies.....</b>	<b>390</b>
<b>IX. Manuscript listing.....</b>	<b>397</b>
<b>X. Acknowledgements.....</b>	<b>403</b>

## Preface

This thesis investigates auditory sensorimotor integration. We define sensorimotor integration as (1) critical involvement of motor systems in sensory processing (sensory-to-motor), and (2) critical involvement of sensory systems in motor processes (motor-to-sensory). A crucial question is to what extent this involvement is really critical. This question has been addressed in some depth but only relating to one of the two aspects of sensorimotor integration, namely motor involvement in sensory processing. Experimental studies dealing with this question reveal that the motor system is activated in action perception, in a way that is congruent (somatotopic) with the performed actions. The motor system “mirrors” the perceived action. Similarly, after musical stroke rehabilitation, patients show motor activation while passively listening to melodies. This result is taken to mean that somehow, the coupling between perception and action is responsible for clinical improvement. However, the criticality of action-perception coupling is challenged by previous findings that the congruency with the performed action can be altered. These views suggest that motor activation during perception merely reflects that the brain has learned to associate certain motor activations with certain sensory activations. This argues against critical involvement.

However, this question of criticality that is amply investigated in motor activation during perception is hardly asked for the sensory-to-motor aspect. That is, it remains largely unclear whether the sensory systems are actually critical to motor processes. This, then, is the topic of the current thesis. We will focus particularly on the timing of motor actions, and investigate whether sensory systems crucially contribute to its learning.

The hypothesis underlying this thesis is that how sensory systems are involved, and whether they are critical, may well depend on the population of participants that we study. For example, musicians, who have been exposed to many thousands of hours of practice, may have so tightly linked the auditory outcome to motor actions that they no longer need it to be physically present. Stroke patients who suffer from motor deficits, on the contrary,

may perhaps use the auditory feedback, but in a different way than healthy individuals.

We will address these questions in the following way. In **part I** we will make more precise what we understand by sensorimotor processing, arguing that scientific progress will be hampered if we fail to make it precise. We investigate the two aspects of sensorimotor integration introduced above in more detail. Understanding the question of criticality of motor involvement during perception will require us to review a substantial portion of the literature on action and speech perception, before turning to music, which is the main topic studied presently. Next, we turn to three populations in which we study auditorimotor integration. First, **part II** investigates the question whether auditory feedback can be used in motor learning in healthy, musical novices. Based on its results, we investigate in **part III** how auditory feedback is used in non-musician stroke patients with motor deficits who are re-learning motor skills. Finally, in **part IV** we ask how auditory feedback is used by expert musical performers. To answer this question, we develop methods of analysis that investigate timing irregularities. For this, we take scale playing as a parallel to the regularity learning used in **parts II** and **III**.

## **I. Sensorimotor integration: a conceptual map**

“[M]a bisogna cercare di capire, lavorando di fantasia, e dimenticare quel che si sa in modo che l'immaginazione possa vagabondare libera [...]”

- Alessandro Baricco, *Oceano Mare*

## I.1. Introduction

---

### 1.1 Aim

Sensorimotor integration is a term that is used a lot<sup>1</sup> but different authors mean such different things by it that it is difficult to understand what the term really means. No review paper exists that combines even a substantial portion of the various fragmented subsets of literature about sensorimotor integration.

The aim of this introduction chapter is to fill this void: to provide a systematic survey of the literature on sensorimotor integration. Its goal is to be an overview and a conceptual framework in which the various parts of the literature can be placed. We will also try to be critical, pointing out over-generalisations and missing pieces of evidence. Such insights open the road for many more future experiments than could possibly be performed in the framework of this thesis.

Since the surveyed literature is broad, the present review could never be exhaustive. The aim is to be synthetic, but not superficial. As a result, we may sometimes point the reader to more specific reviews that are more exhaustive for a particular area.

---

1 A search in PubMed ([www.pubmed.org](http://www.pubmed.org)) on 28 August 2013 revealed that 15,371 scientific references use the word “sensorimotor” in their title or abstract. The keyword “sensorimotor integration” yielded more than a thousand publications.

## 1.2 Disclaimer: finding a balance between unity and diversity

How should we approach such a heterogeneous subject as sensorimotor integration? The aim of science is to identify and describe underlying laws that govern seemingly different phenomena. That is, the hope is to be able to reduce the overwhelming multitude of observables to a small class of laws that allow us to understand what is crucial about their properties. For example, an important breakthrough in science was achieved when it was realised that planetary movements were governed by exactly the same laws that caused an apple, once released from a tree, to drop to the ground. In this way, two very different phenomena (movement of the planet and the falling of the apple) could be united by postulating a single law (gravity). Consider the following alternative example. Just after the sun sets, a particular planet often becomes visible. In the early days, it was thought to be a star, and therefore was called the *Evening star*. Similarly, just before the sun rises, a point of light would be visible, that by analogy was called the *Morning star*. With the advent of more advanced astronomy, it was realised that both were actually the same physical object, namely the planet Venus. Again, two distinct phenomena had been united into a single cause.

The scientific intuition that drives us to always look for unifying causes and reducing phenomena to the minimum of governing laws is certainly a most valid one (see for example Ockham's razor). However, it can sometimes be misleading. Note that in the last example Venus had been referred to as a *star*. We now know that it is not a star but a planet. For a long time, planets and stars were thought to be the same. What concerned scientists of the day was that “planetary stars” move relative to stars, whereas the stars remain fixed relative to each other. Why would two instances of the same phenomena exhibit such different behaviours? The insight that allowed the pieces of the puzzle to fall into place was that actually the “planetary stars” were *not* stars. They were fundamentally different. Against the current of scientific thought that tried to unify the phenomena, it was recognised that planets had to be distinguished from stars. A similar case was that of atoms, which were initially thought to be all the same like refrigerators from a mass-production factory. It would be their differential arrangement that gave rise to the various chemical substances. However, once it was recognised that there were fundamentally different atoms, the scientific field could advance. There may be a parallel to this in the current scientific

view of sensorimotor integration. For example, sensory suppression (see section 5.1.b) or feedback control (see section 6.1.c) have both been called sensorimotor integration. However, a closer look at what these phenomena are shows that they are strikingly different, both in how present computational models implement them and in the underlying neural substrates (Hain et al., 2000).

This review will try to approach sensorimotor integration drawing inspiration from these great scientific developments of the past. We will therefore not try to unify the various phenomena that are referred to by the term *sensorimotor integration* at all costs. As a matter of fact, our scientific progress may be hindered by considering two phenomena identical just because they have been given the same name. Instead, we will present our framework and place the various phenomena in it, inviting the reader to make up his or her own mind. In order to facilitate this process, we will be careful in defining the phenomena under question. The reader may feel that this is tedious and overly scholarly. However, in doing so we provide a closer aim for critics of this work at which to shoot, which is after all how scientific understanding advances most. Furthermore, our hope is that it may prime greater clarity in our thinking on the matter.

## I.2. Preliminaries

---

Some may say that in the brain everything is connected to everything and therefore one should not divide the brain into subsystems such as the auditory or the motor system. Perhaps, to some extent, this is true. However, the brain connection network is astoundingly *small-world*, meaning that generally it is much more clustered than would be expected if all connections were random. At the same time any node in the network can be reached from any other node through a small number of steps (Bullmore and Sporns, 2009). This means that there is also a great deal of structure in the brain networks that will allow us to identify more or less clearly demarcated networks that are responsible for particular kinds of processing. In particular, we will have to start out by defining more clearly what the motor areas are in order for us to be able to meaningfully talk about its integration with auditory areas. Furthermore, we will introduce in this section the various methodologies

that are used to measure motor activity or interfere with motor system processing.

## 2.1 The human motor system – a brief overview

The motor system is defined as all parts of the organism that enable it to move. In what follows, we will pay some particular attention to the ways the brain controls muscles to effect movement, but also to the ways in which the brain receives information about the state of the muscles. Generally, information that streams out of the brain into the effectors is referred to as *efferent* information; the information that is carried back to the brain is *afferent*.

### 2.1.a Motor control areas – a brief outline

The primate's motor system is a complex and multifaceted system implicating nearly every structure in the organism. However, several key areas can be identified and their roles clarified. In this section, we will briefly review these areas and point out particular properties that are relevant to our discussion later on. Clearly, this survey can never hope to be exhaustive, and therefore the interested reader is referred to various textbooks (Rosenbaum, 2009; Schmidt and Lee, 1988).

#### i. Muscles and spinal cord

Limb movements are achieved through muscle fibres contracting in response to electrical stimulation from motor neurons. Muscle fibres are grouped into *motor units* which are innervated by a single motor neuron. The muscle bodies also contain sensory receptors known as *muscle spindles* that the central nervous system uses to detect the present muscle length (also often referred to as *stretch receptors*). Furthermore, the *tendons*, the tissue that attaches the muscles to the bones, contain receptors of their own. These so-called Golgi tendon organs' neural response is roughly opposite to that of the sensory cells in the muscles themselves: tendon receptors fire during muscle contraction, whereas muscle spindles fire when muscles lengthen. Muscle activity can be picked up on the surface of the body through a method called *electromyography* (EMG). Upon muscle contraction, the muscle cells generate electric fields which are registered through EMG and

generally taken as an objective measure for the onset and amplitude of muscular activity.

## ii. Cerebellum

The *cerebellum* is one of the lowest level motor areas in the human brain. It has been implicated in regulation of muscle tone, that is, controlling of the stiffness of limbs (Matthews and Rushworth, 1958). Patients suffering from cerebellar lesions often show problems producing movement sequences such as alternating hand position between palm up and palm down (dysdiadochokinesis) (Holmes, 1939). Similarly, these patients often exhibit problems in timing alternating contractions of agonist and antagonist muscles (Hallett et al., 1975), yielding behaviour that is similar to that of healthy individuals having taken alcohol. Furthermore, the cerebellum appears to play an important role in the learning of skilled movements. For example, firing rates in cerebellar cells were associated with motor adaptation (Gilbert and Thach, 1977). Finally, the cerebellum has been suggested to provide the neural substrate for *internal models* (which will be discussed later in this chapter) (Wolpert et al., 1998).

## iii. Basal Ganglia

The *basal ganglia* (or *basal nuclei*) form a set of subcortical structures that have been implicated in motor control. The structures are particularly densely connected to structures in the cortex (which will be discussed below) and to the thalamus. Amongst other functions, the basal ganglia have been suggested to play a crucial role as a timekeeper in a variety of motor tasks, providing an essential pulse that makes movement possible, especially when this movement requires a regular rhythm such as walking (Brittain and Brown, 2013). Parkinson's Disease (PD) is associated with widespread loss of basal ganglia functioning and consequently patients' movements suffer from “freezing” or tremor.

## iv. (Primary) Motor cortex

The cortical structure that is most directly associated with movement is called the primary motor cortex. When humans initiate movements, cortical activation is spread across the entire brain around 1.5 seconds prior to movement initiation. However, closer to the movement itself, around 50msec prior to EMG-detectable activity the cortical activation

focuses on the primary motor cortex (Deecke et al., 1969). This suggests that the primary motor cortex is more involved in the immediate execution of motor acts and not their planning. Such interpretation is further strengthened by the observation that activating the neurons in the motor cortex through *trans-cranial magnetic stimulation* (TMS) directly causes muscles to twitch. In TMS, a rapidly changing magnetic field is created just above the participant's skull. This field gives rise, through induction, to weak electric currents in the brain tissue. In this way it can cause activity in targeted brain areas. To measure TMS-induced muscle twitches one could simply visually observe the movements. A more subtle way is to record electric muscular activity from the skin surface through EMG. Recording EMG in response to TMS stimulation of motor areas (typically the primary motor cortex) yields measurable electric activity in muscles that are referred to as *motor-evoked potentials* (MEPs).

**v. Premotor cortex**

The primary motor cortex projects predominantly to distal muscles (i.e., far from the torso) such as the fingers. The premotor cortex is a cortical structure that lies just anterior to the primary motor cortex, but, contrary to the primary motor cortex, it projects mainly to proximal muscles (i.e., close to the body center). This has led to the hypothesis that the premotor cortex is involved in postural control and orienting of the body (Wiesendanger, 1981). Electrophysiological studies suggest that premotor cortex may be specifically involved in action anticipation. Monkeys that were trained to respond to a warning light that indicated movement direction, but wait until a go signal occurred, showed premotor activity between the warning light and the go signal, but not after (Weinrich and Wise, 1982).

**vi. Supplementary Motor Area**

The supplementary motor area (SMA) has been implicated in the planning of motor sequences as well as bimanual coordination (Brinkman, 1984; Kermadi, Y. Liu, A. Tempini E.M. Ro, 1997). The area is located dorsal to the premotor cortex. The area typically shows widespread activity 1 second before movement initiation (Deecke et al., 1969) and is active during imagination of a movement as well as actual movement execution, contrary to the

primary motor cortex, which is active only in the latter case (Roland et al., 1980).

### 2.1.b Control Strategies

Now that we have presented a map of the core motor areas in the brain, one might wonder how these areas are used to control movements. The brain's movements are not random but clearly (and often successfully) achieve some goal. In what way is this control over movement achieved? In brief, control strategies that are of interest to our discussion are *feedforward* and *feedback* models, which we will now discuss in more detail.

#### i. Feedforward

Suppose that we wish to generate a grasping or reaching movement to a particular target, for example to a cup placed on the table. One of the simplest ways to control a system (in our case, the *motor* system) is to pre-program what it should do at what point in time. For example, if we want to make our arm reach a cup, the motor system could pre-program which muscles to contract at exactly what time. However, such a system would be very rigid: it would not be able to grasp cups correctly if our muscles were tired, or if our arms grew heavier over time, or if we were injured.

#### ii. Feedback

Another way to make sure that our arm actually reaches the target is to keep our eyes open and continuously track the positions of the target and our arm. The brain can draw a mental arrow from our current hand position to the position of the cup. This arrow encodes the direction and length that our arm should travel to reach the goal. This mental arrow is referred to as the *error signal*, since it indicates whether we are off the target and by how much (Fine and Thoroughman, 2007; Robinson et al., 2003; Sober and Brainard, 2009; Wei and Körding, 2009). It constitutes the feedback and in this case it is *visual* in nature, but one can imagine a similar scenario in which we use tactile information instead. The problem about relying on feedback is that the feedback is (a) sometimes not there (because for example our arm may be blocking the view of the cup for a few instants), (b) delayed (because by the time the visual information reaches the brain and is understood, the state of the world has changed), and finally (c) noisy.

A third class of models called State-Feedback Control models address these problems.

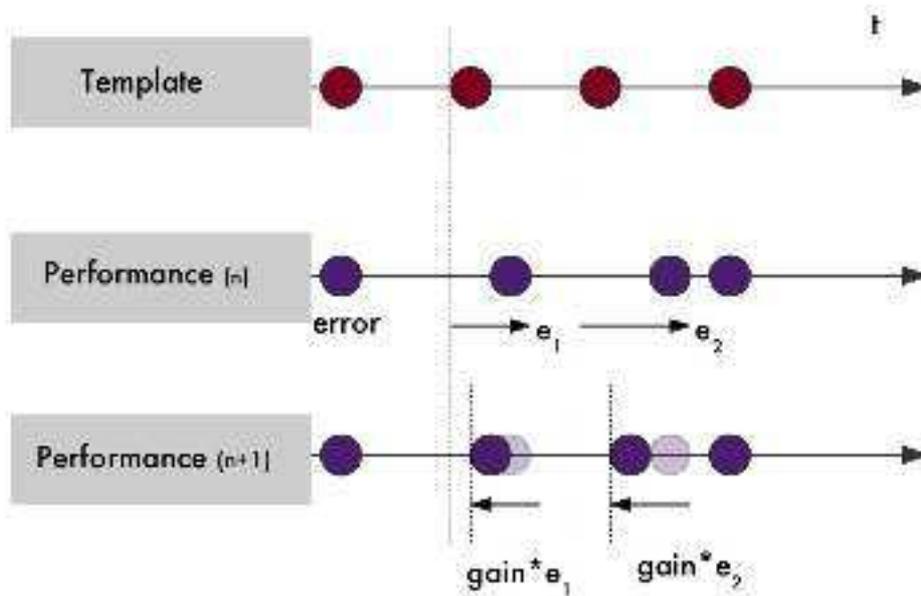


Illustration 1: Feedback-based learning

### iii. State Feedback Control (SFC models)

State Feedback Control (SFC) models (Shadmehr and Krakauer, 2008; Jacobs, 1974) solve the problem of feedback models' reliance on continuous, immediate (not delayed), non-noisy sensory input by constructing an internal model (state) and updating it in real-time using copies of the motor commands that are issued. The model can then be used to generate a feedback signal that is used instead of the sensory signal. Generally, one distinguishes two kinds of such internal forward models (Wolpert et al., 1995): (I) models that make predictions about the state of the motor apparatus (on the basis of the previous state apparatus and the motor commands issued since), and (II) models that make predictions regarding the sensory consequences (again on the basis of the given state of the motor system).

In sum, control strategies have been suggested to function either feedforward (not

using sensory information) or feedback.

### I.3. Towards a definition of sensorimotor integration

---

#### 3.1 Sensorimotor integration

The term sensorimotor integration has a wide variety of meanings and is used by different researchers in different ways. In this section, we will make a start towards a definition that is precise and yet broad enough to encapsulate its most common uses.

**Sensorimotor integration (I)** in its first sense is defined as referring to the phenomenon in which the perceptual and motor systems are both involved in both observation and action. That is, firstly, the perceptual system is involved in production, and secondly, the motor system is involved in perception. We have generalised this from the definition introduced by Hickok et al. (2011).

First of all it is important to point out that these ideas presuppose that we have a clear definition of what the perceptual and motor systems are and that they are distinct. For example, why would we call a certain brain area an *auditory* processing area? We would arrive at this designation based on (a) a series of experiments in which we assume participants are engaging in auditory processing and in which we find this area to be activated; and furthermore (b) an additional series of experiments in which participants are assumed to be not engaging in auditory processing and in which these areas are *not* activated (see Zatorre (2007) for similar ideas). An analogous procedure would allow us to establish which brain areas to consider *motor* processing areas. Here we run into a contradiction: finding that certain areas are involved in both perception and action means that neither will fulfil the criteria for being an auditory or motor processing area in the first place. A consequence of a perceptual-motor overlap is that, strictly speaking, the systems in question can only be designated as sensorimotor. Indeed this is the position some

researchers have taken (Prinz, 1997).

**Sensorimotor integration (II)** in its second sense (distinct from the sense in which it is used above), is defined as the process of combining sensory with motor information so that the organism can effectively interact with its environment. In this case, *integration* refers to the computational processing step that combines (possibly with several transformations, see section 6.1.b) the sensory and motor efferent streams. Note that it is trivial that sensory and motor systems are involved in the process, because after all the task at hand was to combine information from them. Rather, sensorimotor integration in this second sense points to the fact that it is amazing that the brain is able to do it in the first place. For example, Wolpert et al. (1995) use the term in this way: how does the brain combine efference copies of motor commands with perception (proprioception) in arm movements to create a single estimate of the arm position?

Furthermore, notice that sensorimotor integration in its second sense is a prerequisite for most instances of sensorimotor integration in the first sense. The reason for this is that if motor systems are involved in sensory processing, or vice versa, then there must be some way in which information is exchanged between the two systems. The sensory information must somehow be transformed into motor information, or vice versa, or both must be transformed into an intermediate representation. So although the two phenomena are distinct, they are related, and the second sense refers to a more specific process that is required for the first.

## 3.2 Aim

In this chapter, we will review the literature on sensorimotor integration. The main questions are what the role of the motor system in perception is (**section I.5**) and what the role of the perceptual system in motor functioning is (**section I.6**). For both of these sections, we will investigate three phenomena in this order: action in general (**sections 5.2, 6.1 and 6.2**), speech (**sections 5.3 and 6.3**) and music (**sections 5.4 and 6.4**).

The main question will be to what extent the involvement of one system really contributes to processing in the other. This question lies at the heart of the debate in motor theories of perception: to what extent is motor involvement *critical*, and to what extent is it

merely *associative*? The converse of this question is *to what extent is sensory feedback critically used in motor learning*? We will show that the latter has been investigated much less than the former; that is, much effort has gone into addressing this question from the perspective of motor involvement in perception, but very little has gone into addressing it from the perspective of perceptual involvement in motor processes. The present thesis addresses this gap, in particular, we will examine the use of auditory feedback in the learning of movement timing.

The reason we do not restrict ourselves to auditory-motor interactions alone, (n)or the extent that they occur during music performance, is that we believe there is much to gain from putting our discussion in this wider context. All too often speech perception scientists do not talk to speech production scientists (Hickok et al., 2011) and neuroscientists do not interact with psychologists. The aim of this review is to show connections between the lines of thought pursued in these different fields, and to construct a coherent picture of what sensorimotor integration is about.

Why will we discuss speech even though this thesis does not experimentally investigate speech? There are two important reasons. First, speech has been the focus of a great deal of scientific inquiry and therefore a much greater body of evidence and models are available. We will take a great deal of inspiration from the reflections of speech scientists. Second, there is a natural affinity between music and speech. Both are carried by an auditory signal and the motor implementation in creating this auditory signal is not essential to the function of the auditory signal. For example, for as far as the speech content is concerned, it does not matter whether one says: “The leaf is green” whilst standing on one's head, chewing a cookie or while taking a bath. In some way, we recognise that the speech content is similar (even though the acoustic representation of the speech could be severely distorted by the cookie). In other words, the auditory signal is what is controlled by the brain instead of the particular motor implementation (Nieto-Castanon et al., 2005; Kelso et al., 1984). The same appears to hold for music. That being said, recent studies show that concert audiences are greatly influenced by seeing the movements of a performer, and perhaps more than by the sound (Tsay, 2013).

## I.4. Motor theories of cognition

---

In English one can metaphorically say that you *grasp* an idea, playfully suggesting that motor processes are involved in the purely mental act of thinking. Indeed, some theorists postulate that motor systems are involved in all cognition. An example is the *behaviourist* school of thought. According to behaviourism in its most pure form, thought is merely a subtle form of movement. That is, no thought can exist without movement. For example, it was proposed that “there is nothing in the mind that has not been explained in terms of movement” (Pillsbury, 1911, p. 84). A special role seems to be played by the speech apparatus: “thought processes are really motor habits in the larynx” (Watson, 1913, p. 174). Some experimental studies appeared to support this view. For example, when people simply think (which happens to most of us), EMG was found in laryngeal muscles (Sokolov, 1972), indicating that the individuals made speech-like movements that were not strong enough to produce audible speech, so-called inner speech or subvocal speech, see also Oppenheim and Dell (2010).

However, experimental evidence soon discredited this view. In a spectacular experiment a researcher allowed himself to become entirely paralysed (Smith et al., 1947). During this period, he was able to have thoughts and understand what people were saying to him, as he reported when later he recovered his movement abilities. What this shows is that behaviourism in its ultimate form is untenable. However, it is important to note that the fields of neuroscience and cognitive psychology subscribe to behaviourism in a more subtler form. Behaviourism is a response to *dualism*, the position that there exist “mind matter”, a position that was most ardently defended by Descartes. A problem of this position is that it, too, turned out to be untenable, since mind and matter should interact, but if they are separate substances it is impossible for them to do so (Dennett, 1993). As most of our scientific theories, behaviourism arose as a counter-movement against dualism (taking “movement” quite seriously). In its most extreme form, it is not defended anymore. Instead, people opt for a more subtle form, such as the position that the motor system is *crucially* involved in cognition.

Positions that are today realistically held in scientific debate include *embodiment*, which is the idea that the features of the physical body of an individual, (beyond just the brain) play a significant role in cognition (e.g., Thomas, 2013). Again, it tends to be agreed

that its strongest form is untenable, as is the opposite view that the body plays no role whatsoever in cognition (Meteyard et al., 2012). What role would the body play in cognition? The idea is that perception is shaped by action. That is, perception is not a function that can be separated from an action, but rather it comes into being as a result of a confrontation of a (potential) actor with an environment (Gibson, 1977; Greeno, 1994). Interestingly, some patient populations exhibit this action-content of perception quite overtly. When they are presented with objects they immediately grasp and use them even when explicitly told not to or when it serves no purpose (Lhermitte et al., 1986; Lhermitte, 1983). These ideas were later incorporated in an influential paper (O'Regan and Noë, 2001) in which it was argued that perception arises through detailed knowledge of the relationship between sensory and motor events (so-called sensorimotor contingencies) (see also Buhrmann et al., 2013). In similar theoretical accounts, perception evolved in such a way that it is influenced by the costs and benefits of actions. For example, when participants wear a heavy backpack the slope of a ramp seems higher (Proffitt, 2006). The reason is that walking up the ramp became considerably more costly when participants were wearing a heavier load. Arguably it is evolutionarily beneficial for our perception to reflect this directly rather than requiring the brain to draw this conclusion separately.

Notice that there is a subtle difference between theories of embodiment and motor theories of cognition, in that the former consider any feature of the body (such as one's posture or external loads on the body in forms of backpacks) influences cognition, whereas the latter holds it is only one feature of the body, namely movement. However, for our present purposes we will content ourselves to point out that these scientific views exist, and furthermore treat them as equivalent in so far as they postulate motor involvement in cognition. The interested reader may consult recent reviews on embodiment (Proffitt, 2006), and specifically Davis et al. (2012) for a discussion including music cognition.

## **I.5. Motor involvement in perception**

---

In this section, we discuss the first of the two sides of sensorimotor integration, namely the idea that the motor system is involved in sensory processes. We present a

synthesis of a broad range of crosstalk between human sensory and motor systems. Therefore, let us first take a moment to present the abstract structure that will be a recurrent theme in the experimental paradigms we report on. This structure is depicted in illustration 1. We have a motor level and a sensory (or, specifically, auditory) level, and the two are related in some specified way (auditory-motor mapping). For example, whenever we make a particular articulatory speech gesture, a sound will occur, and this sound is not random, but related to the particular movement we made. Similarly, when we press a key on a piano, we will hear a piano tone (if the piano is not broken). Which tone will sound depends on our movement. That is, events happen in the motor domain (movements) and in the sensory domain (sounds or visual events), and these events are related pairwise. The kinds of movements and the kinds of auditory events will differ in the different sections of this chapter. When we talk about speech, the individual movements will be the generation of the individual phonemes (or segments) and the acoustic events will be their sounds, and sequences of them may be words or phrases. When we talk about music, the acoustic events will be notes, and a sequence of them will be a melody.

An individual may have knowledge about the motor and auditory streams individually. For example, she may know how a melody continues when she hears the first few notes, or she can guess the ending of a sentence when she hears the beginning. This is knowledge of the structure on the *perceptual* level and it is represented by horizontal arrows in the illustration. Alternatively, a person may have knowledge about how a sequence of movements continues when he knows how it begins, such as a dancer who may be able to continue a choreography after seeing the first few moves, or monkeys imagining seeing the experimenter picking up an occluded peanut. These examples show that it is possible to have some form of knowledge on of the sequential content of the level of perception or movement separately.

Sensorimotor integration relates to the way the brain represents the association between the movement and perception streams. This integration can happen on two levels, and it is important to distinguish these two. First of all, there is integration on the level of *individual* sensory and motor events. For example, in the case of a piano a person may know that striking one key will give one sound, and striking another key will give another sound. This is a kind of mapping that *associates* movements to perceptual events (an

auditory-motor mapping). Secondly, there can be integration of *sequences* of events. For example, hearing a particular melody, a person may be able to deduce how it was played. Similarly, in the case of speech, it has been suggested that short sequences of speech sounds such as syllables would be stored in an integrated, sensorimotor fashion (Eckers et al., 2013). This second kind of integration may happen without the first: for example, upon hearing a melody she has been trained to play, a person may know how it is played, but given any single note she might not be able to associate it with a keystroke (which is, for instance, the case in Lahav et al., 2007).

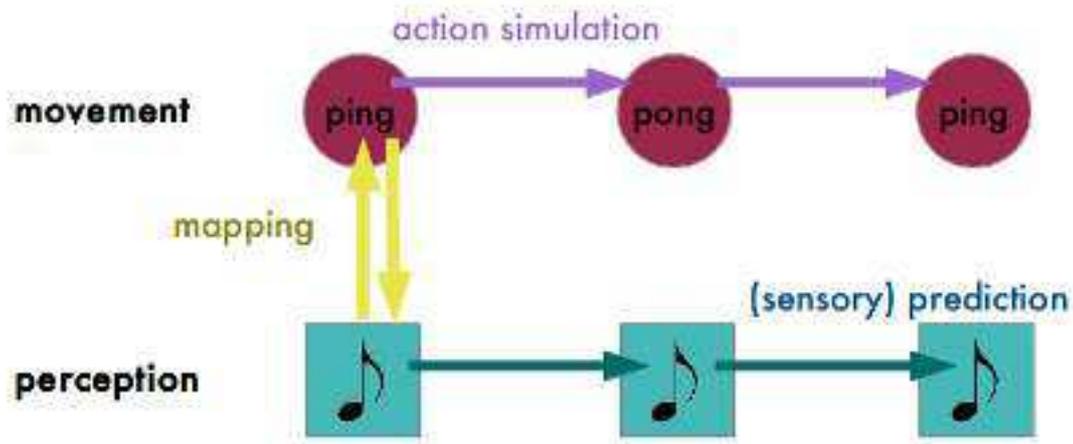


Illustration 2: Element vs. sequence integration.

## 5.1 Movement making perception possible

In order to perceive, we need to move. In order to be able to smell, our lungs need to pump air through our nostrils. In order to be able to feel textures, our fingers need to move (Gibson, 1962). Similarly, in robotics attempting to implement perceptive systems, the idea that one needs movement in order to implement perception is called *active sensing* (Schroeder et al., 2010; Saig et al., 2012). In this section we will point to two phenomena that implicate motor activity in human perception, focusing on the visual and auditory

domains. We take the example of microsaccades in vision. Then we discuss a second example of motion-perception crosstalk, which is the mechanism by which organisms are able to distinguish between their own movement and movement in the environment.

### 5.1.a Microsaccades (vision)

Our eyes bring objects into the fovea through fast jumping movements referred to as saccades. During saccades visual sensitivity is drastically decreased (we are nearly blind), but during the inter-saccade interval perception becomes possible. It was long thought that the eyes remain motionless during these intervals. In fact, during “fixation” periods the eyes slowly (a) drift, (b) undergo superimposed tremor, and, most interestingly, (c) perform small rapid *microsaccades* at the approximate rate of once or twice per second (Martinez-Conde et al., 2013; Rolfs, 2009).

What is the purpose of these small eye movements? Upon their discovery, equipment was designed that would counteract these saccades by displacing the visual image in the same direction as the microsaccade. A remarkable discovery was made: under this manipulation, the visual image fades in a matter of seconds (*perceptual fading*; Ditchburn and Ginsborg, 1952). Troxler, a Swiss philosopher, had proposed as early as 1804 that objects in the visual periphery that do not move disappear (the *Troxler effect*). The microsaccade research in the fifties appeared to confirm that this may be possible even for objects appearing in the fovea. This led to the interpretation that the visual system is optimised to process changes in visual stimulation at the cost of losing from view scenes that stay the same. As a result, once changes in the visual field are absent during fixation, perception fades and artificial changes are induced through microsaccades to bring back visual experience. However, later this interpretation of the microsaccades' function was challenged as observations were reported that only irregular, continuous movement (that is, drift-like motion) counteracted perceptual fading (Gerrits and Vendrik, 1970). Other authors suggested that microsaccades were the oculomotor system's “busy work”, resulting from the fact that fixating for longer periods of time is an unnatural thing to do (Rolfs, 2009).

Nevertheless, over the last few decades, consensus appears to favour the interpretation that some form of change is necessary for sustained visual experience. The oculomotor system achieves this through small movements such as drift or microsaccades,

or a combination thereof, as well as movements of other body parts such as the head (Martinez-Conde et al., 2013; McCamy et al., 2012).

### 5.1.b Motor-induced suppression

When organisms perceive, they need to separate the sensory information originating from their own movements from those originating from changes in the environment. For example, sensitivity to tactile stimuli is reduced during active hand movements because they would overwhelm the senses (Demairé et al., 1984). Motor-induced suppression (MIS) (also sometimes referred to as sensory attenuation) is the phenomenon that the sensory consequences of an organism's own motor acts are perceived as less strong than physically identical stimuli that do not result from the organism's actions. Put differently, sensory suppression distinguishes internal from external sensory inflow. For example, auditory cortex neurons in the primate brains respond less strongly to self-produced vocalisations than to those same vocalisations played back to them afterwards (Eliades and Wang, 2003; Müller-Preuss and Ploog, 1981). During eye movements, the brain suppresses visual input (Bridgeman et al., 1975). Similarly, the chirping of crickets is so loud that they would deafen themselves instantly. However, at every chirp, the sensory system has been warned and the sensory consequences are suppressed (Poulet and Hedwig, 2003). Furthermore, auditory feedback due to chewing in human beings is suppressed (Rosenzweig and Leiman, 1982) and the cortical auditory response to our own phonation is attenuated relative to a played back version (Houde et al., 2002). No suppression occurs when auditory feedback is replaced by white noise (Numminen and Curio, 1999), suggesting that the suppression is not merely a forward process from the motor system to the sensory system, but also modulated by sensory input itself. The brain recognises when it is listening to itself and therefore tunes itself down, but does not suppress sensory input when it realises that it is listening to something else.

How is this attenuation possible? Clearly the sensory systems have been informed of *when* the organism performs *what* movement. This communication is achieved through an efference copy (Holst and Mittelstaedt, 1950), which is a copy of the motor command as dispatched to the motor system, delivered to the sensory (or other) cortical systems. Using the efference copy, the brain is able to predict the nature and the timing of the sensory

consequence of the motor action in question. This predicted sensory consequence is usually referred to as the corollary discharge (Sperry, 1950) (for a review, see Crapse and Sommer, 2008). The organism can then adjust the sensory sensitivity accordingly. As we have seen in section 2.1.b, a translation of intended motor actions to sensory consequences is called a *forward model*. This explains why we cannot tickle ourselves: since our brain is able to predict the tickle sensation before it actually happens, the surprise element that is present when other people tickle us is absent (Blakemore et al., 2000; Weiskrantz et al., 1971). In other words, self-produced tickles are suppressed.

An important consequence of the theory above is that the sensory attenuation should be susceptible to learning. When at first we do not know what the sensory consequences of an act will be, the prediction is that no sensory attenuation will occur. For example, when we are learning to play the piano we have no idea which movements will generate which tones and what qualities of the movement (such as the speed of the keystroke) influence the sound. Indeed, arbitrary movement-to-sound associations can be created and result in the corresponding sensory suppression. It was shown that sensory attenuation can occur and be measured as an attenuated EEG response to self-triggered tones (Schafer and Marcus, 1973). Auditory events such as tones typically evoke a N1 response that is measured at the scalp level using EEG. The amplitude of this response is found to be reduced when the tone is self-triggered, and this happens even after a short training block of 60 trials (Martikainen et al., 2005). That is, once the brain learns to associate tones with keystrokes, it starts to suppress the tones but only just after keystrokes. The attenuation is stronger for an identical stimulus when the participants believe they are the agent relative to when they are made to (erroneously) believe they are not (Desantis et al., 2012). Similarly, the suppression appears independent of attentional modulation (Timm et al., 2013), local in time (Chen et al., 2012b) and dependent on the predictability of the tone effect (Bäss et al., 2008; Hughes et al., 2013). Motor-induced suppression happens in speech signals as well, and is less when the speech feedback is distorted, such as by shifting its pitch (Heinks-Maldonado et al., 2005; Chen et al., 2013b). Sensory suppression appears to also occur for jointly produced actions, although to a lesser extent than individually produced ones (Loehr, 2013).

What is the influence of the learned timing of the tone relative to the movement? In

a recent study (Aliu et al., 2009) participants repeatedly pressed a button at self-paced intervals of about two seconds. The keystroke triggered a tone, and the authors found using MEG brain scans that the N1 response in the auditory cortex was suppressed relative to a control block where tones were presented without the participant pressing a key. This suppression did not occur immediately, but only after a training block. This reflects the fact that the participants needed to train to be able to associate their motor act with the sensory consequence (tone) in order to predict the occurrence of the latter. Furthermore, there was no suppression for a delayed tone. Interestingly, in a different experiment participants were trained in blocks where the tone always came a 100msec delay after the participant's keystroke. The authors found motor-induced suppression still did arise and only in this case also generalised to zero-delay test tones. Furthermore, suppression generalised to alteration of the motor act (switching hands) or the sensory consequence (hearing a tone of a different frequency). Unfortunately, as the authors point out, the two conditions (training with 100msec delay, training with 0msec delay) cannot be directly compared since in the 0msec-delay-experiment, a shorter training block was used.

Recently, the theory that the brain needed to predict sensory effects of its actions for suppression to occur was challenged. In another study (Horváth et al., 2012), participants were exposed to tones occurring at unpredictable intervals and were invited to also press a button whenever they wanted. Although in this experimental setup no causality existed between the keystrokes and the tones, nevertheless the authors found the cortical response to the tones in the temporal vicinity of the keystroke to be suppressed. This finding challenges the theory that sensory suppression is due to the brain's prediction of the sensory consequences of its actions.

It is interesting to see how sensory suppression parallels *temporal binding*, which is the phenomenon by which motor actions and their results appear (to the actor) to be closer in time than they really are (Haggard et al., 2002). However, to our knowledge no studies have directly shown a relationship between the amount of sensory suppression and temporal binding. A simple and elegant study could measure both quantities for the same participants and correlate them.

## 5.2 Action perception

First, we discuss motor involvement in action perception in general. In later sections we will turn to perception of speech (section 5.3) and music (section 5.4), which are special forms of action observation.

### 5.2.a Mirror neurons

Among the most fascinating recent discoveries is that of mirror neurons: double-duty neurons that fire during both action observation and action execution but not in other cases. This class of neurons has been suggested to play a pivotal role in sensorimotor integration, although this idea is still debated. In this section, we will first present the basic phenomenon (sections i-iii) and then investigate more deeply what role mirror neurons might or might not play in sensorimotor integration (sections iv-vi). Mirror neurons were initially discovered in visual action observation. However, in light of our future discussion of speech and music, we will at the end of this section discuss auditory mirror neurons (section vii).

#### i. The phenomenon: double-duty neurons

Gallese et al. (1996) reported finding single neuronal cells in the macaque monkey's frontal cortex that fired both when observing an action as well as performing the action. The neurons were not triggered when seeing either the agent or the object manipulated in isolation. One might argue that the phenomenon reflects the fact that seeing an action *reminds* us of performing it; that is, these neurons might simply reflect motor activity that becomes *associated* to observation. Gallese et al. (1996) argued against this interpretation, showing no M1 activity and no EMG activity on the muscles in question. Furthermore, mirror neurons did not fire when only the object or only agent was presented nor when the action is mimicked without an object.

#### ii. Brain areas containing mirror neurons

Initially, 92 mirror neurons were found among 532 non-mirror neurons in the macaque monkey's inferior frontal gyrus area F5 (Gallese et al., 1996). The F5 area is the

motor area that controls hand and mouth movement and has been studied in considerable depth. It contains “canonical neurons” that fire in response to particular objects (Rizzolatti and Craighero, 2004). Furthermore, the area is thought to contain a “motor vocabulary,” that is, representations of motor actions in the form of neurons that fire when the monkey performs specific actions (such as grasping-with-the-mouth or grasping-with-the-hand) (Rizzolatti et al., 1988).

Subsequently, mirror neurons were also identified in the inferior parietal lobule (IPL) (Fogassi et al., 2005) as well as the monkey's primary motor cortex (M1) (Tkach et al., 2007). The latter finding is interesting, since the original study (Gallese et al., 1996) had not found M1 facilitation and used this as an argument against *associative* motor activity.

The human area BA44 (Broca's area) was initially considered to be the homologue (functional counterpart) of the macaque F5. However, subsequently it was thought that the human ventral premotor cortex (vPMC, Brodmann Area 6) is the more likely candidate (Morin and Grèzes, 2008). This is interesting in light of the finding that the human vPMC is involved in processing sequences of abstract stimuli, regardless of whether they are biologically meaningful (Schubotz and Cramon, 2004). This matter will be discussed in more detail in section 5.3.b.

### **iii. Precisely when are mirror neurons activated?**

The firing pattern of mirror neurons (i.e. the situations in which they fire and the situations in which they do not fire) is surprisingly specific. Mirror neurons only fire for goal-directed (*transitive*) actions such as grasping an object (Gallese et al., 1996) but not for objectless actions such as moving one's arm. A subset of the mirror neurons are triggered even when the object that is grasped or manipulated is not directly seen but only inferred (Umiltà et al., 2001). Mirror neurons also appear specific to whether the action is performed with a tool or not. Ferrari et al. (2005) find neurons in the monkey's F5 that respond even stronger when observing an action performed with a tool than when the monkey performs the action herself with the hand. However, other studies found no or reduced mirror firing during observation of actions using tools (di Pellegrino et al., 1992; Rizzolatti et al., 1996).

**iv. Action understanding**

A debated question is what role mirror neurons play in cognition. Firstly, they have been proposed to be functional in action understanding (Rizzolatti and Craighero, 2004). However, the debate has henceforth been hampered by lack of clarity about what is meant by action understanding (Hickok, 2009). Gallese et al. (1996) defined action understanding as “the capacity to recognise that an individual is performing an action, to differentiate this action from others analogous to it, and to use this information in order to act appropriately” (p. 606). Similarly, Rizzolatti et al. (2001) propose that action understanding is “the capacity to achieve the internal description of an action and to use it to organise appropriate future behaviour” (p. 661). More precisely, “[e]ach time an individual sees an action done by another individual, neurons that represent that action are activated in the observer’s premotor cortex. This automatically induced, motor representation of the observed action corresponds to that which is spontaneously generated during active action and whose outcome is known to the acting individual. Thus, the mirror system transforms visual information into knowledge.” (Rizzolatti and Craighero, 2004, p.172). Nelissen et al., (2005) state that “[a] mere visual representation [of an action], without involvement of the motor system, provides a description of the visible aspects of the movement of the agent, but does not give information critical for understanding action semantics, i.e., what the action is about, what its goal is, and how it is related to other actions” (p. 332). Now notice that the concept of the goal of an action is equally elusive to definition. Is the goal of an action (a) an “abstract” act, or (b) an object (e.g. when you grasp), or (c) a hypothetical state of affairs in the world? Each of these seems to insufficiently capture the notion and be associated with interpretative difficulties of their own (Uithol et al., 2011). The notion of action understanding has subsequently been generalised in humans to include speech perception (Rizzolatti and Arbib, 1998; Wilson et al., 2004).

**v. Adaptive task-dependent sensory-motor associations**

As a result of problems associated with the action understanding interpretation of mirror neuron functioning, more cautious interpretations have been advanced that consider the mirror neurons to represent *adaptive task-dependent sensory-motor associations* (Hickok, 2009) (see also Mahon and Caramazza (2008) for a related critique of embodied cognition).

Recall that initial interpretations had called upon the absence of “overt” motor activity to argue that mirror neuron activity is not simply associative. Tkach et al. (2007) then found that, contrary to previous reports, the macaque M1 also contains mirror neurons. This undermines the idea that no motor activity is involved in action perception (Gallese et al., 1996). Instead, the finding suggests that the mirror neurons may after all be simply a sign of the motor system associating certain actions to the action percept (for example through Hebbian learning). That is, during observation, the motor system is “getting ready” to move, without quite moving yet (the so-called set-related response; Wise and Mauritz, 1985). Similar interpretations have been advanced towards mirror neuron's role in social functioning (Keysers and Perrett, 2004).

#### **vi. Other proposed roles for mirror neurons**

Mirror neurons are alternatively thought to be the neural basis for imitation (Iacoboni et al., 1999; Rizzolatti and Craighero, 2004), but this is rendered less credible if one takes into account that the macaque monkey species does not imitate (Hickok, 2009). Additionally, Gallese and Goldman (1998) argue that the mirror neuron may be the neural substrate for mind-reading (i.e. creating a representation of people's intention). Fogassi et al. (2005) similarly suggest that mirror neurons serve to identify the intentions of the actor.

#### **vii. Mirror neurons responding to auditory observation**

The mirror neurons originally discovered responded to visual observation of actions. Is there anything special about the visual modality or do mirror neurons also exist for other modalities? Kohler et al. (2002) identified visual mirror neurons in macaque monkeys and found that about 15% of them responded also to action-related sounds such as ripping paper or cracking a peanut. This suggests that this particular subset of mirror neurons encodes action in a modality-unspecific way (Keysers et al., 2003). It remains important to note that the large majority of mirror neurons do not have this property but instead are specific to one modality, such as vision.

Furthermore, swamp sparrows are found to possess auditory mirror neurons (Prather et al., 2008) that are song-specific and fire during both listening to and singing that song. However, we should note that these recordings were made in a "singing back" paradigm

and therefore the nature of the task may have given rise to the exhibited activations. Subsequently, these neurons have been referred to as “Echo neurons” (Zatorre et al., 2007). To our knowledge, auditory mirror neurons (i.e. single auditory mirror neurons, rather than auditory mirror networks) have not been identified in humans.

## 5.2.b Mirror networks

### i. Distinguishing mirror neurons and mirror networks

Single-cell recordings in healthy humans are problematic for obvious ethical reasons. Therefore the counterpart of the monkey studies reported in the previous section could long not be performed in humans. However, fMRI brain scans revealed neural networks in the human cortex that have properties comparable to mirror neurons. That is, these networks activate during action performance and action observation (Gazzola and Keysers, 2009). Functional MRI does not have sufficient resolution to pick up single neuron activations and therefore these results do not show conclusively that humans have mirror neurons. If neurons involved in *either* action observation *or* action performance are spatially sufficiently intermingled, one will obtain the same global network activations for observation and execution without there actually being single “mirror” neurons involved in both (Illustration 3). In other words, we may find *mirror networks* with single-duty observation and single-duty execution neurons heaped together, without any double-duty *mirror neurons*. Note that it would be an interesting finding in itself that observation and execution are so closely intermingled, but it does not suffice to demonstrate mirror neurons in human beings.

For this reason, it remained unclear whether mirror *neurons* exist in humans. In 2009, Hickok published a passionate attack against existing assumptions, arguing that no evidence showed mirror neurons to exist in humans (Hickok, 2009). One year later, single-cell recordings of activity in the human medial frontal and temporal cortex found evidence of mirror neurons that fired both for action observation and action execution (Mukamel et al., 2010). Interestingly, the study also revealed “anti-mirror” neurons that were active during action performance but inhibited during action observation.

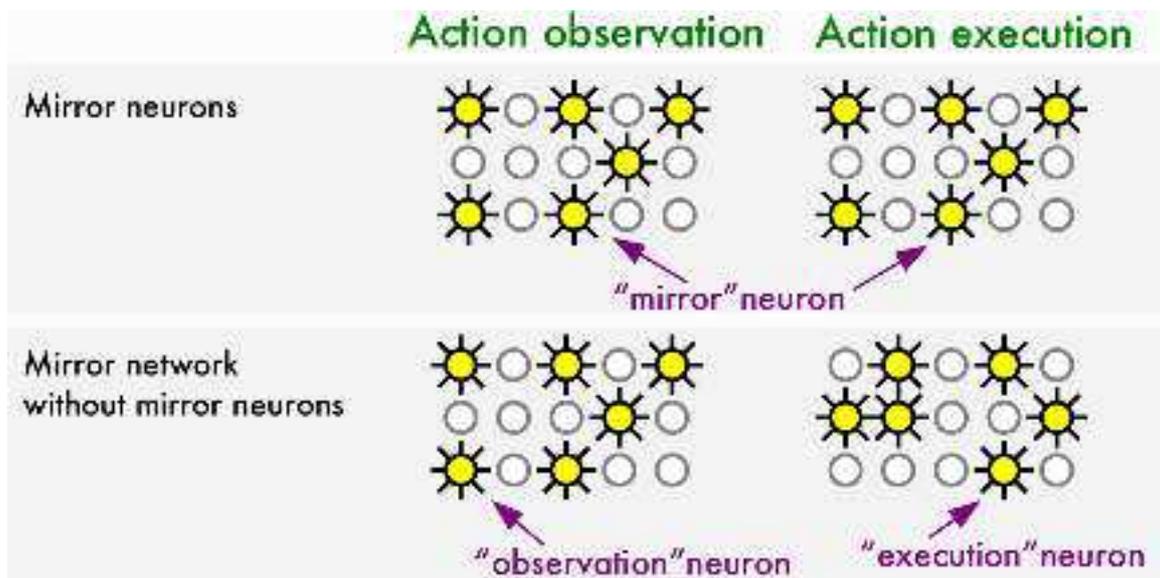


Illustration 3: Explanation of the difference between mirror networks and mirror neurons.

## ii. Visual action observation

In humans, evidence for *mirror networks* had been established long before evidence for mirror neurons. Fadiga, Fogassi, Pavesi, & Rizzolatti (1995) applied TMS to the motor cortex of humans observing motor actions. They found the resulting MEP to be significantly increased relative to a control condition in which their participants did not observe motor actions. This finding is in line with the interpretation of mirror neurons as a *set-related response* (see section 5.2.a.v) which is further strengthened by the finding that motor activity during observation is enhanced when participants know they later will need to report specific properties of the action (Schuch et al., 2010).

The mirror network, like mirror neurons, activates specifically during observation of biological motion. For example, investigating whether the observed actor needed to be a biological entity, Tai et al. (2004) found no mirror network activity when observing a robot perform an action, whereas (Gazzola et al., 2007) did find premotor activation, leaving the issue of whether the observed agent needs to be biological still contended. Related to this, Stevens et al. (2000) show participants slide-shows of movements at various speeds. When the slides were shown at high-speed, participants considered the motion impossible because

the experimental setup implies the limbs would have had to cross. At lower speeds, the movement was perceived as biologically possible, as there is time for the limbs to pass alongside one another. Motor activation was more pronounced during perception of possible motion than during the perception of biologically impossible motion. In line with this interpretation, a different study revealed that rTMS over the human premotor cortex disrupts processing of action-related pictures (Urgesi et al., 2007), which the authors interpreted as an involvement of the human motor system in action processing (although this interpretation is disputed; Hickok, 2009). Similarly, rTMS over the vPMc impairs visual perception of movements but only of those that are biologically possible (Candidi et al., 2008). Similarly, Kilner, Paulignan, & Blakemore (2003) show that participants exhibit less motor activity when observing a robot do the action, suggesting an influence of characteristic biological movement profiles. Perhaps this is due to the fact that biological motion tends to have minimum jerk (Kilner et al., 2007). In line with this reasoning, motor activation is more pronounced when observing slide-shows of stills depicting more effortful actions (Proverbio et al., 2009). Kinematic measurements have furthermore revealed that the observation-induced motor activation can facilitate simultaneous overt movements themselves during a fairly brief time window (Ménoret et al., 2013).

The motor activation during action perception appears to be more closely linked to processing of the goal than of the action itself. Indeed, premotor activity was enhanced when participants were asked to report the goal rather than the effector of observed actions (Hyman et al., 2013).

What is the influence of expertise on this mirror network activity? Several studies have suggested that experience performing the motor action in question subsequently enhances mirror motor activity when observing the action (Calvo-Merino et al., 2010, 2005). Indeed, dancers' self-rated capacity to perform dance movements correlated with ventral premotor activity during passive observation (Cross et al., 2006). Furthermore, motor activity was enhanced for movements belonging to those dancers' repertoires relative to movements they only observe (because they belong to the opposite gender) (Calvo-Merino et al., 2006). However, it has furthermore been suggested that simply observational learning may explain this effect, since dance lovers show motor activation when observing other dancers, even when the latter are doing movements that they cannot do (Jola et al., 2012).

For a review of (pre)motor activity during action observation, see (Morin and Grèzes, 2008).

### iii. **Auditory action observation**

The previously discussed studies have almost exclusively investigated visual observation of actions. In the light of our interest in auditory-motor integration, we will here discuss particularly auditory (“echo”) mirror networks.

Indeed, mirror networks parallel to those reported for visual action observation are found for the auditory modality. For example, corticospinal excitability of hand-muscles was increased when participants heard hand action sounds relative to when they heard leg action sounds or non-action sounds (Aziz-Zadeh et al., 2004). Similarly, distinct facilitation patterns were found for hand actions as opposed to mouth actions (Gazzola et al., 2006). That is, the motor activity appears to be somatotopically related to the action that produced the sound in question. However, this finding appears to be contradicted by other studies in which the motor activation did not depend on the effector (Galati et al., 2008), perhaps due to a difference in experimental paradigm. For a more detailed discussion of auditory somatotopy in mirror activations, see Aglioti and Pazzaglia (2010).

Similarly, premotor activity is often reported in response to affective non-speech sounds such as laughter or screams (Meyer et al., 2005), and the activation appears proportional to the extent to which the stimulus is arousing (Warren et al., 2006).

### **5.2.c Direct matching hypothesis and motor resonance**

The evidence reviewed thus far indicates a mirror network, probably grounded to a large extent in mirror neurons, that *mirrors* the observed action by mapping it onto motor representations (Aglioti and Pazzaglia, 2010). This mapping appears to be to some extent independent of the modality (visual or auditory) of perception and nevertheless highly specific to the action in question (as revealed by the somatotopy of the mapping).

Put simply, the idea is that the motor system of an observer somehow matches the motor system of the observed actor. Motor resonance is the phenomenon that when one observes an action, the motor system of the observer “resonates” with the motor system of the actor (and mirror neurons may be responsible for this). For example, participants

observing a drummer show a similar time-course of M1 activation to the drummers themselves (Caetano et al., 2007). Contrary to what the term suggests “resonance” is not taken to imply that the observer's motor system engages in synchronised action with the motor system of the executor of the action. Rather, resonance implies that the activity in both systems is somehow comparable (Uithol et al., 2011). Motor resonance is thought to be closely linked to social interaction. Indeed, individuals who have just engaged in social interaction show enhanced motor-evoked potentials during human actor observation relative to those who have not, suggesting that social interaction *primes* mirror networks (Hogeveen and Obhi, 2012). Typically researchers distinguish between *intra-personal* resonance, in which the motor system resonates with the perceptive system, and *inter-personal* resonance, in which the motor systems of observer and observed directly resonate (in a way that is mediated by the perceptual system).

What would be the role of this resonance? One proposed mechanism that became known as the Direct Matching Hypothesis is that the motor representation used to execute an action is “matched” to the perceptual representation of the action observation, and this process enables the observer to understand the action (see section 5.2.a for a definition of action understanding and its problems). One piece of evidence for this idea comes from a study showing that eye movements (which are crucial in performing visually guided movements) are strikingly similar when a person observes another performing an action and when the person executes the action herself (Flanagan and Johansson, 2003). Furthermore, purely motoric (non-visual) learning is shown to improve subsequent visual recognition of the movements in question (Casile and Giese, 2006). The result is difficult to explain without resorting to the observer's activation of motor representations “matching” those of the one performing the action. However, critical stances relative to these observations suggest distinct pathways for perception and action (Goodale and Milner, 1992).

#### **5.2.d Plasticity of mirror networks: arbitrariness of the associations**

Is the mirror activation changeable or does it remain the same across one's lifetime? The functioning of the mirror networks is surprisingly pliable. Catmur and colleagues (2007) applied TMS over the motor cortex of human observers who were shown a

slide-show of pictures of a human hand of which either the index or the little finger moved. The MEP measured in the muscle of the corresponding finger in the observer was greater than that of the other finger (which was not seen moving). This is in line with mirror network somatotopy which had been discovered before (section 5.2.b.ii). However, next the experimenters trained their participants to move the *opposite* finger than the one shown moving in the movie (i.e., they were instructed to move the index finger in response to seeing the little finger move on the screen). After this so-called counter-mirror training, the MEP in the finger opposite to the one observed moving was greater than the finger not observed moving. In other words, the motor somatotopy had been completely reversed. Furthermore, mirror (pre-existing) and counter-mirror (experimentally manipulated) motor activation due to action observation have a comparable time course (Cavallo et al., 2013), adding to evidence that they may be largely the same process.

Similarly, Petroni et al. (2010) trained participants to associate actions with neutral stimuli such as colour patches. The fact that the percept of these colour patches later also evokes motor activity shows that the nature of the stimulus modulating the sensorimotor coupling is fairly irrelevant. This, of course, is problematic for an interpretation of motor networks as crucial for perception, because it is difficult to imagine how motor networks contribute to the participant's percept of colour if they were not activated from the start (see also Press et al. (2012) for a similar paradigm with geometric shapes).

The finding that mirror networks are changeable shows that the way the motor system is modulated by action observation is not fixed, but depends on the observer's experience (Catmur et al., 2008; Catmur, 2013). This fact presents an important problem to the interpretation that motor activity reflects processes involved in the understanding of action. The mirror network plasticity has two important consequences: (1) it shows that the associations are fairly recent in an evolutionary sense, not being hard-wired in our mammal brains (see also Gazzola et al. (2006) for a discussion), and (2) it makes it less likely that the motor activation is crucial for *understanding* the action. The reason for the latter point is the following. Imagine that we used one of the paradigms mentioned above to change a participant's associations between finger movements and visual observation of finger movements. Suppose that we found that as a result, the cortical motor somatotopy is altered. That means that observing the same action, participants will show different motor

activation. If one claims that motor activation represents action understanding, one is therefore committed to admit that somehow we have changed our participant's *understanding* of that action.

These considerations led to the formulation of the Associative Sequence Learning (ASL) theory of mirror neuron functioning (Heyes, 2010) (see also Del Giudice et al., 2009). The premise of this theory is that mirror neurons are initially non-specifically linked to a wide variety of sensory neurons. With continued exposure to combinations of sensory and motor information (for example through observation and imitation but also to any self-initiated action that yields sensory consequences) the mirror neurons become more specifically tuned to observation of the motor actions that these neurons code for. That is, through a process of Hebbian learning the brain learns to associate more and more specific actions to specific observations. This theory therefore accounts for the canonical mirror neurons but also the fact that mirror networks are changeable.

We return to these points later in our discussion of the influence of expertise and familiarity in auditory-motor co-activation in music studies (section 5.4.d). It also gave rise to the theory that action-perception circuits implicated in language (section 5.3) arise due to repeated association between actions and their consequences (Pulvermüller and Fadiga, 2010).

To summarise this section (5.2), experimental evidence provides a strong support for the claim that motor systems are activated during perception. Initially, these observations gave rise to the interpretation that its role is *critical*, so that their role is to enable the individual to understand other's actions. However, the pliability of the motor activity speaks against this interpretation, suggesting the motor system might instead be merely collaterally and associatively activated. For a recent discussion of these fascinating issues, see Avenanti et al. (2013). We furthermore point out that even if the motor systems' involvement in perception is not critical, it may still be beneficial.

### 5.3 Speech perception and comprehension

We now turn from action observation in general (discussed in section 5.2) to the observation of a specific type of action, namely speech. The reason why the speech literature is of interest to us is that it is *communicative* whereas many ordinary actions (such as grasping a cup) are not. This makes speech an excellent bridge to our discussion of music in section 5.4. Furthermore, the speech signal is more restrictively defined than action in general and therefore the models discussed in this section are more precise.

With the term speech perception, we will refer to the process by which humans are able to perceive and recognise speech sounds. Comprehension is the process by which the spoken words are combined in sentences and understood. Specifically, we will be interested in the processing steps that are unique to speech as opposed to general auditory processing. The process of speech perception has many parts and these will be discussed in the coming sections. For one, the sounds will have to be recognised as speech sounds, isolated from non-speech sounds, and categorised into phonemes (section 5.3.a). Further, the brain will need to integrate these phonemic sounds into sequences that allow for lexical access and subsequent syntactic (section 5.3.b) and semantic processing (section 5.3.c). Finally, there are higher-level processes involved in ecological conversational settings (section 5.3.d). In each of these processes, the motor system has been suggested to be involved, but in each case this is for different reasons. Therefore, we will discuss each of these processes in turn and discuss the alleged motor involvement separately.

#### 5.3.a Perception of speech sounds

First, let us focus on the earliest step: that of phonemic perception. The same speech sounds exhibit a remarkable variability in their acoustic realisation, and phonemes that sound the same may be very different in a spectrogram representation. That is, there is no clear one-to-one mapping between perceptual representations of speech sounds and their acoustic realisation. So how is the brain able to perceive and categorise speech sounds? The motor theory of speech perception (Lieberman and Mattingly, 1985) addresses this question. The fundamental premise of this theory was that during passive speech perception, listeners track the articulatory movements that produce the speech sounds they hear. That

is, “perceiving speech is perceiving vocal tract gestures” (Liberman and Mattingly, 1985). Technically, the acoustic signal is thought to contain cues which allow the listeners to postulate possible articulatory actions, which are then internally simulated and finally compared against the actual percept. Recognising the movements that generated particular speech sounds would allow the listener to deduce which speech sounds the speaker made. The idea is that this step allows the speaker to bypass a purely acoustic analysis of the speech sounds (subject to the problem described above) but of course required theorists to posit the necessary machinery for auditory-to-motor transformations. That is, somehow the speaker needs to know which auditory events map to which articulatory gestures. Therefore speech theorists have criticised the motor theory of speech perception for merely having moved the problem one level up (see Galantucci et al. (2006) for a recent review of the motor theory of speech perception and refer to Schwartz et al. (2012) for variations to the basic theory formulated in the *Perception-for-Action-Control Theory*). A crucial prediction of this theory is that the motor system will be active during (passive) speech perception. Indeed, there is a wide variety of data supporting activation of the motor system during speech perception. Notice, however, that the observation of motor activity is a necessary but not sufficient condition for the motor theory of speech perception (Bever and Poeppel, 2010). A second prediction of the motor theory of speech perception is that disruption of motor activity should lead to a disruption of perception. Thirdly, patients with lesions in motor systems are predicted to exhibit perceptual deficits. In the following sections we will discuss these predictions in turn.

#### **i. Motor co-activation**

It is a common finding that hearing speech sounds increases activation in motor and premotor areas relative to baseline silent conditions. This is reflected in increased MEPs of the muscles around the mouth during auditory or visual speech observation (Fadiga et al., 2002; Watkins et al., 2003). Similarly, fMRI scans of participants perceiving speech reveals increased motor activity relative to baseline (Wilson and Iacoboni, 2006; Wilson et al., 2004) in a way that is somatotopic (Pulvermüller et al., 2006). That is, when hearing sounds produced by the lips (such as labial sounds), the corresponding (lip) area of the motor cortex of the listener is activated. In brain scans, activity is most commonly found in the

premotor cortex. The motor activation seems to appear only for particular task conditions, such as when participants are asked to classify speech sounds (Hickok and Poeppel, 2007) or if the speech signal is distorted (Oblaser et al., 2005). Similarly, Davis and Johnsrude (2003) found stronger premotor cortex activity in response to degraded speech than to normal speech. Greater motor activity may be elicited by rare words (the *lexical frequency effect*, Roy et al., 2008). For an overview of motor activation in response to speech and speech-like sounds, see (Scott et al., 2009).

## ii. Disruption of motor activity in healthy individuals

A similar prediction of motor involvement in speech perception is that disruption of motor system should impair speech perception. Indeed, disrupting premotor cortex functioning using TMS impairs speech discrimination performance in noise (Meister et al., 2007). Furthermore, disruption of motor activity of particular articulatory effectors (such as lips or the tongue) specifically affects perception of the sounds made by those articulators (D'Ausilio et al., 2009b). That is, the disruption is *somatotopic*, in line with the neuroimaging studies showing that motor activation during perception is somatotopic. In addition, disruption of parts of the motor system not involved in speech articulation (such as hand areas) has no influence on speech sound categorisation (Möttönen et al., 2013). However, Lotto et al. (2009) argue that TMS studies do not show a noticeable difference in motor activation during speech perception and non-speech control sounds (Fadiga et al., 2002; Watkins et al., 2003), which implies that the role of motor systems in speech processing is questionable.

A second line of research shows the intricate links between motor learning and speech perception. In a classic study (Nasir and Ostry, 2009) participants were instructed to pronounce some 300 words while their jaw was forced into protruded position. Over the course of the experiment, participants learned to compensate up to 50% for this displacement (as is commonly found in motor learning paradigms). The novel finding was that those participants who show this adaptation also show a perceptual shift, perceiving /had/ more as /heard/ in a subsequent listening test. This perceptual shift furthermore correlated with the amount of motor adaptation. The result shows that disruption of the motor system affects perceptual functioning. Furthermore, alterations in sensory systems

appear only after active but not after passive motor learning (Ostry et al., 2010), thus ruling out explanations based purely on somatosensory experience.

### iii. Lesion studies

A further prediction of a motor theory of speech perception is that lesions in motor areas should affect speech perception. Studies of patients with language deficits due to brain damage seem to provide a more nuanced picture, suggesting that motor lesions do not necessarily lead to perceptual difficulties. *Expressive aphasia* (also called *non-fluent aphasia*, *Broca's aphasia* or *agrammatic aphasia*) is a condition in which patients show normal auditory speech perception and comprehension but impairments in speech production. This being said, expressive aphasics have been reported to also exhibit perceptual deficits in a phoneme categorisation task (Basso et al., 1977). However, these deficits did not correlate with their speech perception and comprehension capacities. These findings suggest that the motor system influences or modulates speech perception but is not directly used during speech perception. That is, damage to the motor system does not necessarily lead to impairment in speech perception, as would be predicted when we postulate that the motor system is critically involved in speech perception. *Receptive aphasia* (also called *fluent aphasia* or *Wernicke's aphasia*) is the opposite condition, characterised by a loss in speech perception and comprehension, but normal, fluent speech production (although the produced speech is typically senseless) (Crinion et al., 2006). This condition seems to invalidate the claim that the perceptual system is crucial for speech production (which is discussed in section I.6). For a more detailed discussion of the lesion data, see Scott et al., (2009).

### iv. Development

Further evidence against critical involvement of motor areas in speech perception comes from the observation of children growing up with conditions such as dysarthria or anarthria, which cause severely impaired or no speech production. These children nevertheless develop normal speech comprehension (Bishop, 1988). Also, in normal infants, speech perception is usually fully developed before they engage in speech production (Werker and Yeung, 2005).

To summarise this section (5.3.a), a wide variety of empirical data supports activation of motor areas during speech perception. The activation appears largely to conform to mirror network activity (as previously discussed) in the sense that it is somatotopic. However, evidence from lesion and development studies indicates that the motor system is not crucial for speech perception. Consequently, the current prevailing opinion is that the motor theory of speech perception in its strongest formulation is probably false, leading researchers to propose that the motor influence is *modulatory* rather than *critical* (Toni et al., 2008). This is in line with accounts of mirror network functioning such as those discussed in section 5.2.

### 5.3.b Syntactic processing

We now turn to motor involvement in syntactic processing. The role motor systems are thought to play in syntactic processing are different from those proposed in speech sound perception or general action observation. The idea that syntactic processing (or in general, operations on abstract sequences) might rely on cortical motor circuitry is based on the structural similarities and proximities of Broca's area (a usual locus of language-related activity) and premotor areas (Binkofski and Buccino, 2004). The question of whether language abilities are based on general cognitive capacities (e.g. *cognitive grammar*, Langacker, 1987) or are a distinct capacity, independent of other cognitive faculties (Hauser et al., 2002), is still subject of debate. Some of the researchers defending the former position postulate that “syntax is represented and processed in systems that regulate motor output, including speech production.” (Scott et al., 2009, p.298). In other words, motor areas that are responsible for sequential processing are evolutionarily older circuits developed for gestures or other sequential motor acts (Fogassi and Ferrari, 2007; Glenberg and Gallese, 2012). Put differently, linguistic syntax computations are implemented in existing structure-building infrastructure in the brain (Dominey, 1997).

To make this precise, *local structure computation* (such as building simple word compounds) would be located in more posterior premotor areas (Fiebach and Schubotz, 2006; Friederici, 2006) and interact with Broca's area, which would implement higher-order syntactic processing (such as moving syntactic constituents). This is in line with studies

implicating the ventral premotor cortex in processing of abstract structures (Schubotz and Cramon, 2004; Schubotz et al., 2003). Similar support for premotor contributions to syntactic processing comes from observations of impaired finite state artificial grammar learning (but, surprisingly, intact phrase-structure grammar learning) in patients with ventral premotor lesions (Opitz and Kotz, 2012). Furthermore, mutations in the alleged “language-gene” FOXP2 affect the formation of auditory-motor associations (Kurt et al., 2012), suggesting a common neural basis for auditory-motor associations and language processing. However, other studies criticise this interpretation based on the claim that language functions are predominantly left-lateralised whereas sensorimotor circuitry is bilaterally distributed (Aziz-Zadeh et al., 2006).

### 5.3.c Semantic processing

The semantic analysis of perceived speech is the process by which the brain is able to understand the meanings of the words and combine them to understand the meaning of the perceived utterance as a whole. In this process, too, motor involvement has been postulated (Pulvermüller et al., 2006). It is proposed that to understand an action word, the listener exhibits motor activation corresponding to its meaning (Feldman and Narayanan, 2004) in a somatotopic way (Hauk et al., 2004; Kemmerer et al., 2008).

These positions have been supported by experimental data. First, processing of action sentences describing actions of the mouth, hand or leg is shown by neuroimaging studies to somatotopically activate (among other areas) the premotor cortex (Tettamanti et al., 2005). This effect appears to persist (though somewhat attenuated) when the action words are used in a non-literal fashion (e.g. *to grasp an idea*) (Sakreida et al., 2013; Schuil and Zwaan, 2013). A second prediction is that perception of action words should interfere with tool processing and indeed it does (Creem and Proffitt, 2001). Third, more abstractly, motion words appear to influence perceptual discrimination acuity for dot motion (Pavan et al., 2013). Quite surprisingly, it has been suggested that whereas positive sentences with action words enhances motor excitability, perceiving negative sentences inhibits the corresponding muscles (Foroni and Semin, 2013).

Research investigating whether this motor co-activation is *critical* to the understanding of the words is sparse (but see Papeo et al., 2013). Disruption of the primary

motor cortex during action word processing had no noticeable effects (Papeo et al., 2009). Similarly, after learning to associate particular actions to novel words, passive perception of those new words yield only limited evidence for motor co-activation (Fargier et al., 2012). That is, there is a great potential for experimental investigation to clarify whether this semantics-induced motor activation is related to the other kinds of sensorimotor integration presented so far in this survey.

### 5.3.d Conversational interaction

Finally, we discuss the highest computational level of speech perception, which is processing that relates to conversational interaction. The position defended by Scott et al. (2009) is that motor activation during speech perception reflects a conversational attitude by the listener. In conversation, speakers take turns in speaking and the pauses between these turns are astonishingly short, even when speakers cannot see each other and are not acquainted (De Ruiter et al., 2006). This suggests that even during listening, the participant is getting ready to jump in and start speaking, for example to help the interlocutor finish his or her sentence. Passive listening may not be so passive after all. It is argued that the role of the motor system in speech perception is a sign of conversational participation, or creating a common perceptuo-motor ground (Sato et al., 2013): usually when we hear speech, we are about to talk back. That is, motor activation could be compared to a set-related response as was suggested for mirror neurons (see section 5.2.a). This account of motor involvement in speech perception is particularly interesting in light of our upcoming discussion about music, as it has been proposed that motor activity during music perception might equally be explained as resulting from the wish to participate in (e.g. tap along with or dance to) the music.

At this point, one may wonder how the somatotopic motor activation may be explained as being “participatory” motor activity. Recall that upon hearing labial or dental phonemes, the corresponding part of the motor system is implicated (D’Ausilio et al., 2009b). Why would this be, if, during normal conversation, discussants are not making the same sounds at the same time? The proposed explanation is called *convergence*. Convergence is a phenomenon by which participants in a conversation start, gradually and mostly unconsciously, to move similarly. For example, speakers show a tendency to breathe

in synchrony (Scott et al., 2009). Similarly, choral singers synchronise their heartbeats (Vickhoff et al., 2013). Therefore, since speakers' movements will tend to match, matching motor somatotopy in the speaker and listener follows as a prediction of this theory.

### **5.3.e Anatomical substrates**

How would this coupling between the auditory and motor systems be physically implemented in the neural structures of the brain? Speech theorists have developed several models of speech processing. These models are more general than our current interests as they are aimed to explain how the various steps involved in speech processing are implemented. For our present purposes, it suffices to investigate those steps of these neural models that involve the auditory and motor system.

The dual-stream model of speech perception is developed and introduced by Hickok and Poeppel (2004; 2007), generalising previous accounts separating action and perception pathways (Goodale and Milner, 1992). The aim of the model is, roughly speaking, to account for the double dissociations between speech perception and production deficits (see section 5.3.a.iii), whilst still accommodating the interaction between the perception and production systems. The model is based on the work of Wernicke, who had proposed a parallel two-stream speech processing model. The idea of the dual-stream model is that when we hear speech, the auditory information undergoes basic spectrotemporal processing in the parts of the brain that are classically thought of as auditory areas. Next, the information is split into two streams that undergo subsequent parallel processing that is more or less independent. One stream, the *dorsal* stream, maps the sounds to the articulatory system, that is, the motor and premotor regions of the brain. This is the stream that implements the sensory-to-motor mappings (see the introduction to section I.5) and it is probably the cause of most of the motor activation seen during speech sound perception (section 5.3.a). The second stream, the *ventral* stream, maps the sounds to meaning. The ventral stream is most likely what gives rise to the semantic motor activation (section 5.3.c).

### **5.3.f Necessity and role of motor involvement**

At this point, two related questions will need to be addressed. First, is the motor

system *required* for speech perception, as opposed to simply peripherally activated? Second, what is its *role*?

We have observed that the motor system is not *required* for speech perception. This rules out the strong form of a motor theory of speech perception. Instead, a common position is that the motor system *modulates* speech perception (Hickok et al., 2011).

In order to clarify the role motor processes may play, it is important to distinguish between the various processes of speech perception. We provided a starting point for such an account by discussing the contributions motor systems may make to processing of speech sounds (section 5.3.a), syntax (section 5.3.b), semantics (section 5.3.c) or conversational interaction (section 5.3.d). It is likely that for some of these processes, motor involvement is critical, whereas for others it is not. In a similar vein, it has been proposed that perhaps what determines the kind of motor co-activation is not the stimulus itself, but the strategy that the participant is using during perception (Tomasino and Rumiati, 2013).

## 5.4 Music perception

### 5.4.a Why music?

Why use music as an object of study for auditory-motor integration? Existing models of speech perception (discussed in section 5.3) and production (to be discussed in section 6.3) are advanced and there is a great deal of similarity between music and speech. However, music allows us to manipulate the effect of training. This variable is difficult to manipulate in the case of speech. It would, after all, not be ethical to disrupt a child learning such a crucial skill. In music, when they remain within bounds, such manipulations can be acceptable.

Furthermore, human beings are all experts at speech already. We spend our lives babbling away and develop an exceptionally skilled motor apparatus that is able to implement the demands of everyday speech production. In music, people are also experts in music perception (Bigand and Poulin-Charronnat, 2006). After all, we listen to music on the train, as we walk the streets, at home, or when we have dinner in a restaurant. However, only a limited number of people can play a musical instrument. Among those who do, there

is a great variety of skill levels. In the studies presented in this thesis, music production expertise ranges from none at all to up to forty thousand hours of accumulated practice time. This range provides a unique opportunity to study the auditory-motor connections as they come into existence.

Secondly, there are key differences between language and music that have a bearing on the sensorimotor integration discussion. Music does not have semantics in the same way that language has. There is a debate as to whether music has a semantics, which may be based on tension and relaxation. But clearly music does not allow one to express: “My nephew put the bread in the cupboard.” For our purposes, this is a good thing, because it allows us to abstract away from this literal level of meaning and its influences on sensorimotor integration. For example, experiments investigating motor activation during music listening will not have interference from whether the signal semantically expresses a hand or leg action (Hauk et al., 2004) because music does not allow us to express that kind of meaning.

#### **5.4.b Motor systems supporting perceptual memory formation**

It has been suggested that the motor system enhances functioning of several other cognitive faculties, such as perceptual sequence learning. In Deroost and Soetens (2006), participants had to respond to cues appearing in a fixed number of spatial locations by pressing an associated button. Unbeknownst to participants, the cues were not random but followed a fixed sequence, and participants' responses to cues that were part of the sequence became faster relative to responses that did not follow the pattern. This phenomenon and its associated experimental paradigm is known as *serial reaction time* (SRT) (Nissen and Bullemer, 1987). However, if the (perceptual) cue sequence follows a specified pattern then the associated responses (if they are correct) follow this sequence too (Mayr, 1996). By properly controlling for this interference, it is possible to dissociate perceptual from motor learning (Remillard, 2003). The finding is that purely perceptual learning picks up only relatively simple deterministic sequence structure. However, with the addition of active motor involvement learning, increasingly complex patterns became learnable (Deroost and Soetens, 2006). The proposed explanation is that when participants are exposed to a perceptual sequence but concurrently execute a motor sequence, the

perceptual and motor sequences are bound together. Upon recollection, not only the perceptual but also the motor information is reactivated (Nyberg et al., 2000; De Lucia et al., 2009; James and Atwood, 2009; Weisberg et al., 2007), and this enhances memory quality (Barsalou, 1999). This perceptual-motor binding then generalises to multisensory integration as well (Butler et al., 2011; Butler and James, 2013). We would like to point out that it remains unclear whether the active motor involvement is crucial or if passive motor involvement would have the same effect. If so, the effect in question is not due to perceptual-motor binding but entirely perceptual (visual and somatosensory). In order to test this, experimenters could use a control group in which arm movements are reproduced by a robot while the participant remains passive (Wong et al., 2012; Ostry et al., 2010).

The phenomenon of motor-enhanced memory formation has been investigated in more explicitly musical tasks. Indeed, professional musicians' memory for melodies is enhanced if they have not only heard them but actively played them (Brown and Palmer, 2012). Further evidence from neuroimaging corroborates these effects. The cortical activation pattern during auditory exposure to the melody shows enhanced activity in the insular cortex when it is actively learned (Mutschler et al., 2007), suggesting that this cortical region may play a role in learning general action-effect associations. Similarly, Lappe et al. (2008) trained participants to play a four-bar isochronous broken chord progression bimanually over 8 days. The aim was to learn to play more correctly, faster, and more smoothly. In a subsequent perceptual task, participants showed an increase in the mismatch-negativity (MMNm) when hearing a fragment of the trained piece with one changed note as oddball. This increase was absent to a contrasted auditory-only group who only listened to what the other group played during their learning phase. The analogous effect appears when participants are not trained to play a melody but rather tap a rhythmic sequence (Lappe et al., 2011). As a result, the mechanism in question need not be based on a melodic but rather an integrated melodic-temporal representation. The finding of a benefit of (active) motor involvement relative to mere passive perceptual learning is strikingly similar to studies investigating perceptual alterations due to motor activity. For example, Ostry et al. (2010) found changes in somatosensation in participants after they learn to perform a motor task. Crucially, these changes were absent in a control group that did not actively move but was passively moved by the experimenter.

#### **5.4.c Music perception in novices: a focus on rhythm perception**

Motor activation has been reported systematically in studies that investigate passive music listening in musical novices, which is interpreted as evidence for the power of music to move us (Janata et al., 2012). For example, premotor activity is found in listeners anticipating the end of a musical piece (Leaver et al., 2009) or processing a deviant in a chord sequence (Koelsch et al., 2002, 2005). Similarly, motor cortex excitability is enhanced for music that has a strong so-called groove, that is music which is propelling the listener to move, relative to low-groove music or noise (Stupacher et al., 2013). For a review about cortical processing involved in music perception (see Koelsch, 2011).

Rhythm perception appears to play a specific role during music perception and be a key factor for eliciting motor activity in musical novices (Popescu et al., 2004; Grahn and Rowe, 2009; Grahn and Brett, 2007). For example, when participants passively listen to musical rhythms, premotor areas are activated, even in the absence of an instruction to tap along at a later time (Chen et al., 2008). An isochronous (regular) auditory stimulus induced cortical entrainment evident in modulation of the beta amplitude (Fujioka et al., 2012). Conversely, overt motor involvement in tasks is shown to improve timekeeping accuracy in a purely perceptual task (Manning and Schutz, 2013; Iordanescu et al., 2013), probably suggesting that motor activity can sustain a beat when the sounds are absent. The way we interpret the metre of a musical piece is influenced by how we were moved (Phillips-Silver and Trainor, 2005) or how we moved (Phillips-Silver and Trainor, 2007) when we heard it.

These findings, taken together, have suggested an intimate link between sequencing and timing behaviours in music and other domains (Janata and Grafton, 2003), in which the premotor cortex may play a special role (Merchant et al., 2013). The fact that music propels us to move is often cited as evidence for unique connections between auditory and motor areas (Zatorre et al., 2007). However, in the light of the abundance of motor activation during speech or action perception (see sections 5.2 and 5.3), this causality is questionable. After all, when we see others perform action or hear speech, we do not feel compelled to join moving or speaking as we do with music. For example, people will be less inclined to listen to an audiobook on their iPod while running instead of music. This suggests that we may need to look beyond mere cortical motor activation in order to understand what it

really is about music that moves us.

#### **5.4.d Musical experts**

Expert musicians are those who have accumulated many years of instrumental practice. This population's relationship to music is closer to normal healthy individuals' relationship to speech, as we are all speech experts, having practiced our articulators to perfection (although musicians in their lifetime still speak much more than they play their instrument). Therefore, one would expect findings concerning motor activity during speech processing in normal individuals (section 5.3) to be reproduced for music processing in musical experts. As we shall see in this section, this is indeed mostly the case.

##### **i. Motor activation and facilitation during passive listening**

This section discusses motor activation during passive music listening. Experimental evidence for the opposite co-activation (sensory activation during silent performance) also exists and is discussed in section 6.4.a.

Haueisen and Knösche (2001) performed the first neuroimaging study to investigate auditory-motor integration in musicians. The authors compared MEG activation in pianists compared with non-pianists passively listening to familiar melodies. They found that in the 300ms prior to each note, a distinct activation was evident in the contralateral primary motor cortex of pianists compared with nonmusician controls. Furthermore, a differential activation was furthermore found for notes played with the thumb and little finger, showing somatotopic topography of the activation. The fact that the activation occurred prior to each note suggests an anticipation, and is specific (somatotopic), suggesting internal simulation (Jeannerod, 2001) or a form of mental motor imagery (Kosslyn et al., 2001; Lotze and Halsband, 2006) of the particular movements. This interpretation was further supported by fMRI comparing both (silent) music production and (passive) music listening in professional pianists (Bangert et al., 2006). A musicianship-specific distributed cortical network was found comprising premotor areas and supplementary motor areas that were activated in both cases by the professional musicians. The network is triggered not only by auditory perception of music, but also by visual observation of piano playing movements (Haslinger et al., 2005) or the encoding of movement sequences that might

evoke auditory feedback in the future (Pau et al., 2013). Taken together, these observations suggest that this phenomenon is similar to mirror networks (section 5.2.b).

Does listening to music lead to motor facilitation in analogy with speech listening? To answer this question, intra-cortical facilitation (ICF) was measured while amateur pianists listened to familiar and unfamiliar musical pieces. In brief, intra-cortical facilitation is measured by applying a brief subthreshold stimulation burst to the motor cortex (conditioning stimulus) prior to a supra-threshold stimulation. The resulting motor-evoked potential is facilitated when the ISI between the pulses is roughly 8-30msec (Kujirai et al., 1993; Wagle-Shukla et al., 2009). This facilitation increased while listening to the musical material but only for the learned pieces (D'Ausilio et al., 2006). The fact that facilitation was specific to the pieces the participant had learned to play shows that it was not the individual key-to-tone mappings but the activation of sequences of auditory-motor events that primed the motor system. The motor co-activation furthermore depends on the listener's biography. Musicians showed enhanced activity in a distributed cortical network (including the primary motor cortex) when listening to the same musical piece played by their own instrument relative to when it was played by another instrument (Margulis et al., 2009).

Listening can in itself facilitate overt movements in a way that is particular to the movements to be performed. D'Ausilio et al. (2009a) had pianists press piano keys in response to a visual cue in the form of a note in sheet music notation. They examined the priming effect due to a congruent or incongruent tone presented at various points in time relative to the visual cue. They found that early tones (300 to 200 msec prior to the visual cue) facilitated movement when they were congruent. For example, hearing an early D when the pianist was about to play a D facilitated the reaction time to strike the D key. Slightly later (up to 200 msec after the visual cue) there is an inhibition due to incongruent tones. That is, hearing a late F sound when the pianist was about to play a D inhibited the reaction time to strike the D key. The finding crucially shows that pianists' brains were able to map the heard sound to the corresponding key (using an approximately one-to-one mapping of sounds to actions).

Interestingly, the auditory-motor connections are not only transiently established in response to auditory (or, for that matter, motor) activity, but are also enhanced at rest in

musical experts (Luo et al., 2012). Resting state brain scans measure cortical activation when the participant lies passive in the scanner. This has the important advantage over neuroimaging during a particular task, which is that differences in cortical activation between groups of participants cannot be attributed to their differential performance. The finding that auditory-motor connections are enhanced for musicians relative to non-musicians is therefore all the more surprising. As we discussed previously, we are all experts at speech and auditory-motor integration has been implicated in speech. For this reason, it is surprising that Luo and colleagues find no auditory-motor resting connectivity in nonmusicians. Why did their speech expertise not translate into auditory-motor connectivity? This result is difficult to explain if auditory-motor integration in speech and music were fundamentally the same process.

#### **5.4.e Creation of auditory-motor associations: music learning in novices**

The picture emerging from the previous section is that passive music listening recruits motor areas in musical experts, in parallel to the empirical evidence accumulated in action observation and speech perception (sections 5.2 and 5.3). This opens the door to investigating the process by which auditory-motor associations are created in musical novices as they learn to play an instrument.

Pascual-Leone et al. (1995) invited non-musicians to learn a simple 5-tone scale for 2 hours per day over the course of 5 days. Participants became more accurate and less variable in their piano performance over time. These behavioural improvements were accompanied by a dramatic reduction in TMS motor thresholds for motor activity for the trained, but not the untrained, hand. The result shows that musical training effectively induces cortical plasticity, but the link between sensory and motor regions remains unclear. To clarify this question, Bangert and colleagues introduced a “repeat-practice” paradigm in which the nonmusician participants' task was to reproduce on the piano melodies that are presented auditorily (Bangert et al., 2001; Bangert and Altenmüller, 2003). Throughout the experiment, the melodies increased from very simple, few-note melodies to more melodically and rhythmically challenging ones as participants' performance improved. The aim of this paradigm was to disentangle the creation of sensorimotor associations from memory formation. Participants were trained to associate keystrokes with particular

sounds, since their task was to find the motor activity (sequence of keystrokes) that would result in the ever-changing melodies presented to them auditorily. Participants were divided into two groups. In a map-group, the key-to-tone mappings were consistent, in a non-map group they were shuffled in between the practice blocks (i.e., days) but consistent within days. The map-group shows a decrease in errors after the training, but the no-map-group does not, showing that the latter group was not able to learn the task. The map-group furthermore reveals enhanced cortical response (measured through DC-EEG) to auditory probe tones whereas the no-map group did not. However, the spatial resolution of EEG did not allow the authors to pinpoint the cerebral sources of this activity.

A subsequent fMRI study used a similar music learning task and was able to locate co-activation to the premotor region, the intraparietal sulcus and the inferior parietal region (Lahav et al., 2007). Furthermore, the authors found that the motor activation is present only when the auditorily presented melodies are the ones the participants were trained to perform (which mirrors the familiarity effect on motor facilitation in musical experts, see section 5.4.d.i). However, the paradigm used by these authors was different from that used by Bangert and colleagues in that participants learned to play a single musical piece rather than learning a generic sound-to-key mapping skill. Therefore, differences between the findings of various studies could be attributed to differences in experimental paradigms used.

What is the role of motor activation found during music perception? In the absence of overt movement and instructions to move, it is suggested that auditory presentation of melodies evokes mental motor imagery. Indeed, the same motor activations were enhanced when participants were invited to passively imagine performing the corresponding movements (Baumann et al., 2007).

To further clarify the difference between learning auditory-motor sequences and learning learning single auditory-motor associations, the following experiment was performed (Chen et al., 2012a). fMRI data was recorded in participants learning to repeatedly play back a particular melody, interspersed with trials in which participants played back random melodies. It was found, among other things, that the capacity to play back a single note from the trained melody improves, but there was no such improvement for untrained notes. Activation in the superior temporal gyrus (STG) while listening to the

trained melody was lower after training than before training, and during learning PMd and PMv show reduced activity.

Do sensory-motor links learned for one sensory modality generalise to other modalities? To investigate this, Engel et al. (2012) had participants learned to play melodies that were either auditorily or visually presented. Subsequently, participants were scanned in the magnet as they were visually shown those sequences that are learned auditorily. The authors found that the left rolandic operculum exhibited greater activation in response to trained relative to untrained melodies. The region was also involved in silent execution of movement sequences, which makes it plausible that the auditory-visual cross-modal transfer is mediated by motor representations.

For more details about learning music performance, the reader is invited to consult Herholz and Zatorre (2012).

## **I.6. Sensory involvement in action**

---

We now turn to the second aspect of sensorimotor integration, which is defined as critical involvement of the sensory system in motor functioning. We propose to distinguish between sensory *feedback* and *stimulation*. This distinction relies on the connection (or absence thereof) between the sensory and motor event sequences. Imagine the following two situations. In the first case, an individual sits at a piano and strikes its keys, which result in sounds. In the second case, a metronome gives the person a steady beat and she is asked to swing her finger in the air in time with the metronome sounds. The first case describes a sound that forms a *sensory feedback* in the sense that the sound depends on the individual's movement. Its timing depends on when she strikes the key; its pitch depends on which key she strikes; and its loudness depends on how hard she strikes the key. That is, a blindfolded listener would be able to reconstruct, from what he hears, to a large extent, the movements that she has made. In the second example, the (metronome) sound does not depend on her movement. The blindfolded listener, if he cannot hear her finger moving through the air (which is probably the case even for the very acute of hearing), will not be able to reconstruct her movement. For all he knows, she may not have been moving at all.

The point here is that the sensory events do not rely on the motor events at all. This second case will be referred to as sensory stimulation and is contrasted with sensory feedback. Note that we do not argue that the motor system cannot use sensory stimulation (e.g., the metronome) in some aspect of motor control. As a matter of fact, we will see that patient populations can benefit dramatically from sounds that do not depend on their movement (section 6.2.b.ii). However, the way in which the brain uses the sound will be different from a case in which the sound actually tells the brain something about the performed movement.

Another difference between feedback and stimulation is the extent to which movements and sounds are *coupled* in time. In the second case (i.e., sensory stimulation), movements are not automatically synchronised with the metronome clicks. Indeed, this is quite a complex task and humans tend to make systematic errors (Repp, 2005). However, in the first (feedback) case, it is trivial that the movements and the sounds coincided in time (or, if the system contains many delays, at least they will be *time-locked*). This fact is due to the way the piano is built to generate sounds in response to keystrokes. Put more technically, in the case of sensory stimulation, it is surprising (and indeed impressive) when the brain is able to induce a *coupling* between the sensory and motor events. In the case of feedback, this is trivial, because it is inherent in the way the task is set up.

We now turn to a more technical definition of stimulation and feedback. We propose to define **feedback** as a sensory signal that contains relevant and interpretable information about motor production. A very simple kind of feedback could be when one presses a button, one sees a light turn on. This tells the person that she successfully pressed the button. But it does not tell her with what speed she pressed the button. Furthermore, the time the lamp comes on may be quite uncertain and delayed, especially if it is a energy saving lamp for example. Alternatively, one could design an experimental setup (so-called sonification) in which the vertical coordinate of one's finger is mapped to the frequency of a tone that sounds continuously. In this case, when one presses a button, the auditory information will contain much more information. It will allow a blindfolded listener to decide whether she has pressed the button (for example, when the frequency she hears is below a certain value). But contrary to the previous case, she will also know the trajectory of the movement as she moves her finger down.

Furthermore, let us define **sensory stimulation** as (typically predictable) sensory events that do not rely on your movement. Hence, stimulation carries no information about the movement. Simply put, feedback is what sound is in music performance, and stimulation is what the same sound is dance.

Sensory feedback may contain information about several aspects of the movement (for example, spatial or temporal aspects), and each of these pieces of information may have a high or low resolution (information content). In the case of the button press that turns on a light, the sensory feedback contains binary spatial information (on or off), and a more or less precise temporal information. In the case of the sonification paradigm the sensory feedback would have a much higher spatial information content in the signal (depending on how well we are able to hear frequency differences). Note that whether the brain is actually able to use this additional information content in a meaningful way is a different question. Quite possibly the brain will be overwhelmed and not be able to use this information in motor learning.

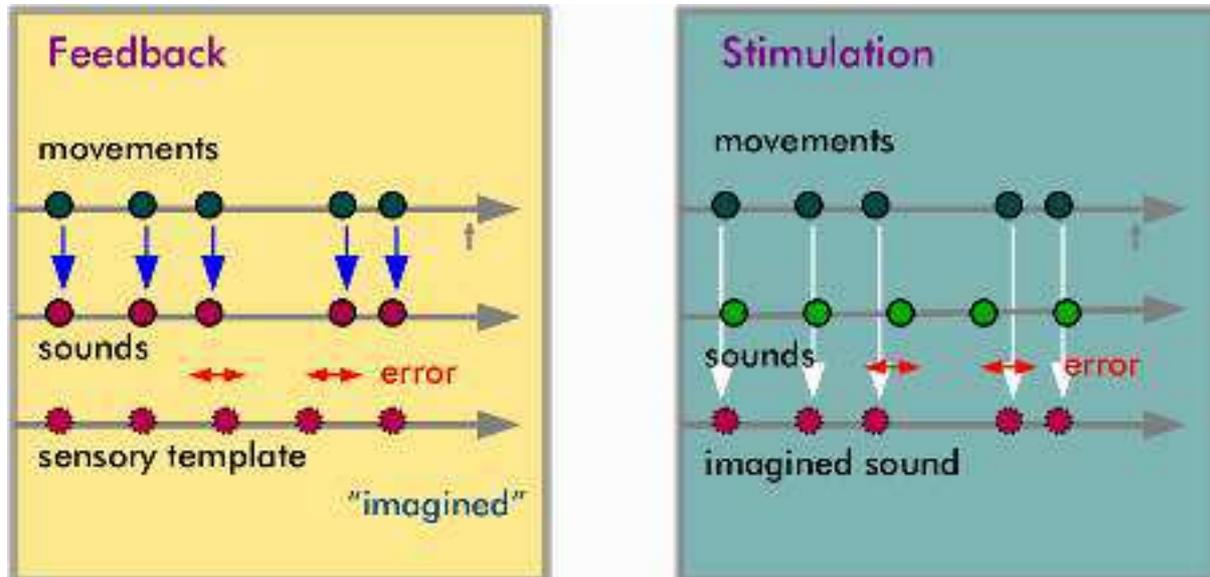


Illustration 4: Sensory feedback vs. sensory stimulation

How would the brain's processing of the sensory and motor events differ for feedback as opposed to stimulation? Imagine that a participant is instructed to generate a regularly timed movement sequence. Illustration 4 shows an example of some motor events (e.g. keystrokes or taps on a table). In a feedback paradigm, the movements generate sounds. If the brain has some template of what a perfectly regular sequence would sound like, it would be able to compare the actually produced sounds with the intended sounds so as to generate an *error signal* (red arrows in illustration 4). The error signal would simply tell the brain that it needed to produce this movement a bit earlier and that movement a bit later, for example. Now if the experimental paradigm were different, for example if one only had a *stimulation* sound such as a metronome, then the brain would obviously not compare the sound to some inner template because that would trivially yield a zero error signal. Instead, in this case the brain would have to try to predict when the sounds would come (given its knowledge about when it sent particular motor commands) (depicted by the white arrows in illustration 4). Then, the brain would be able to construct an error signal between the movement and the sensory consequences as well.

These observations suggest that if the brain cannot generate a regular sensory template, then a feedback setting will not be able to bring any benefit. The reason is that there is no sensory template and therefore no error signal on which to base learning. However, if the brain has a problem predicting sensory consequences of its movements (for example due to malfunctioning internal models), then it will fail in a sensory stimulation setting.

In this section we have categorised studies according to whether the sensory events are feedback or stimulation. However, the sensory events can also be a combination of the two. Repp and Knoblich (2007) instructed their participants to tap along with a metronome (i.e., a stimulation setting). Unknown to them, suddenly the metronome became contingent on their taps (i.e., a feedback setting). These innovative paradigms may enable future investigation of the differences between neural structures involved in feedback and stimulation tasks.

We will now turn to movement-evoked sensation (feedback) and movement-independent sensation (stimulation) in turn.

## 6.1 Sensory feedback in action

In this section, we will discuss the influence of sensory feedback on motor functioning. In particular, we investigate action selection, execution and motor learning.

### 6.1.a Representing action by its sensory consequences (action selection)

Ideomotor theory is a framework for understanding how the brain plans action. The principle idea is that movements are represented by their perceptual consequences. Ideomotor theory was introduced in the psychological discipline by William James. In his *Principles of Psychology* (James, 1918), it reads: “Every representation of a movement awakens in some degree the actual movement which is its object” (p.526). That is, before we are able to move, a sensory image is created for the movement. Subsequently, the movement is executed in such a way as to align with these predicted perceptual consequences: “An anticipatory image of the sensorial consequences of a movement is the only psychic state which introspection lets us discern as the forerunner of our voluntary acts” (p.501). This understanding points at an intimate link between action and perception, namely that (mental) perception is required for action to become possible: “[m]ovement is the natural immediate effect of feeling” (p.527). William James' observations were described introspectively, an approach to psychology that was later abandoned in favour of behaviourism. Ideomotor thinking was later revived, more cautiously formulated, as the ideomotor principle (Greenwald, 1970).

An anecdotal example of how sensory effects are used to represent action is the following. A friend wrote me the following in an e-mail: “[after the rebellion] they one back their freedom.” Now this clearly is not just a spelling mistake. She heard “won” in her head, then typed the homonym “one”. This shows that the sensory (auditory) signal plays a role in the selection of typing movements. This directly supports ideomotor theory, which suggests that we select the action (typing the word) by the sensory consequences (not the keystroke clicks but the sound it produces when read aloud).

Ideomotor thought was further articulated in contemporary thinking as part of the theory of event coding (TEC) framework (Hommel et al., 2001). Rather than being a psychological theory in the usual sense, TEC provides a framework within which to

approach the perception-action interface. The foundation of the framework is, firstly, that physically produced effects and those that are intended are represented in a common domain. Secondly, action and perception are represented in a similar, distributed fashion. Thirdly, the codes that represent events (actions or perceptions) are abstract (distal) instead of concrete (proximal). The fact that action and perception share a coding domain is alternatively referred to as common coding.

Support for the ideomotor theory comes from studies that typically use a two-phase model of action control. In this paradigm, participants undergo a training phase to establish associations between stimuli, responses and effects and then test whether the effect itself can prime the corresponding response (Greenwald, 1970; Kunde et al., 2004). For example, Elsner and Hommel (2001) had their participants press either a left or right key in response to a stimulus. The keystrokes were followed by task-irrelevant low- or high-pitched sounds, respectively. In a test phase, these sounds preceded participants' keystrokes and were shown to influence reaction times and response frequencies. These findings provided support for the interpretation that participants used the tone in the process of response selection. It is important to note that it is *action initiation* (or response selection) that shows a facilitation but not *action execution*. For example, typically the time between the stimulus and response onset is facilitated by response-effect associations, but not the duration of the response or the peak force of the response keystroke (Kunde et al., 2004). This suggests that response-effect associations facilitate response selection, but not response execution (Keller and Koch, 2006b).

The two-phase model of action control is an experimental implementation of the notion that action-effect associations are subject to training. For example, in counter-mirror adaptation studies (see section 5.2.d) serve this exact same purpose.

In a serial-reaction time paradigm, learning is also shown to be improved when tones follow responses relative to when no sounds are present (Ziessler, 1998). The effect requires, however, that there is enough time between the tone and the next stimulus (Hoffmann et al., 2001) and that the same tones are consistently mapped to the same keystrokes (Stöcker et al., 2003).

Associations need not be learned in a two-phase experimental paradigm, however: they can be inherent. For example, strong keystrokes invoking loud sounds and soft

keystrokes invoking soft sounds are learned faster than the other way around (Kunde, 2001), reflecting the fact that humans normally associate strong effects with strong responses. Similarly, when participants are responding to a finger movement video as a cue their responses are faster when the (observed) cue coincides with their response (tapping the finger or lifting the finger) than when it is opposite (Brass et al., 2001). The underlying association between the observed cue (tapping the finger) and the response (tapping the finger) is not induced in the experimental paradigm but present at the outset. Similarly, Keller and Koch (2006a) had participants press a key in response to a colour patch cue. The keys are aligned in a vertical row and each button triggers a tone. When the key-to-tone mapping was aligned with the spatial layout of the buttons (top keys to high tones, bottom keys to low tones) responses were faster than when the mappings are incongruent. In an alternative paradigm where participants were instructed to generate a sequence of keystrokes as regularly in time as possible, compatible key-to-tone mappings again facilitated temporal precision (Keller and Koch, 2006b) and the effect grew with musical experience (Keller and Koch, 2008).

A further line of empirical support for ideomotor theory comes from animal studies revealing the *differential outcome effect*. In this paradigm, animals are taught to give particular responses (behaviours) to particular stimuli. The learning is shown to be enhanced when different responses are reinforced in different ways, for example by giving the animal different kinds of food depending on the response rather than the same kind of food for two responses (Trapold, 1970). The finding is significant in that it shows that when the effect is a reinforcer it is not simply a source of motivation for the animal but an important element in its learning (Estévez et al., 2003).

For a recent review of ideomotor theory, see Shin et al., (2010).

### **6.1.b Sensory guidance of action (action execution)**

In this section, we discuss sensory feedback in the on-line guidance of action. A wealth of empirical studies exist that point to the brain's usage of sensory feedback in maintaining online control of movements. In particular, visual feedback is used in grasping and reaching movements (Thaler and Goodale, 2010; Saunders and Knill, 2003, 2004; Reichenbach et al., 2009; Santello et al., 2002). Studies on deafferented patients exhibiting

impaired motor performance reveal the crucial role of proprioceptive feedback (Gentilucci et al., 1994; Ghez et al., 1995) but also show that visual information might take over in its absence (Stenneken et al., 2006). Auditory feedback appears to also be used by the brain in guiding pointing movements (Boyer et al., 2013). Tactile information is furthermore thought to be important in online error correction in finger tapping (Aschersleben et al., 2001).

A question still subject to some debate is how the brain is able to combine sensory and motor information so as to achieve this on-line feedback-based interaction with the environment. Indeed, this is a different use of the phrase *sensorimotor integration* (see section I.3), where integration refers to the merging of the sensory and motor integration streams (Wolpert et al., 1995) instead of the overlap between the sensory and motor processing systems themselves (as we have defined it above). The question is of considerable interest in vision, where a coordinate transformation will need to be effected between sensory (visual) and motor coordinates as part of this merging of the visual and motor streams. This is a *reference-frame transformation* (RFT) (Blohm and Crawford, 2007; Flash and Sejnowski, 2001), which we will not discuss in further detail here.

### **6.1.c Motor learning through sensory feedback (action learning)**

In the previous section, we investigated how sensory information influences on-line motor performance. However, an independent question is whether sensory feedback may benefit motor learning. Motor learning is the acquisition of a novel motor skill, commonly due to repeated performance of the motor task, resulting in relatively permanent improvement in motor performance (Wolpert et al., 2011; Newell, 1991). For example, when we use a computer keyboard repeatedly, our movements will become faster and more accurate.

It is important to note that manipulations that improve motor performance do not necessarily improve motor learning. In a classic study in the motor learning literature, participants learned to perform a set of different movements. Their performance during the training was better if each of the movements is trained in a separate block rather than when all of the movements are trained interleaved (randomly). But it turns out that when motor performance is measured in a retention test after the short-term effects of the training have faded, learning is greater in the latter case (Shea and Morgan, 1979). The

important consequence is that facilitating movement performance may actually degrade motor learning (Schmidt and Lee, 1988). In other words, the question of sensory involvement in motor *learning* has to be treated separately from the question of sensory involvement in motor *performance*.

Discrepancies were found between motor learning and performance in a sequence learning paradigm. Musicians were trained to tap a sequence of keystrokes and exhibited an improvement in sequence production performance when they received a sound at every keystroke, relative to when they received no sound. This difference was present from the first trial onwards, which means that participants had not yet had the chance to learn anything. Comparing the later trials, the authors found no difference in learning between the sound and the mute group (Conde et al., 2012). The question remains open whether nonmusicians would benefit from sound in similar tasks (and this question will be addressed in this thesis).

It has been proposed that motor learning relies on feedback from various modalities such as vision (Bernier et al., 2006) and proprioception (Wong et al., 2012). Indeed, motor learning itself may be for a large part explained by perceptual learning (Vahdat et al., 2011). If motor performance depends on perception, then an improvement in perceptual functioning only would nevertheless experimentally appear as an improvement in motor functioning, that is, motor learning.

Feedback from different modalities may be used differently by the brain. For example, Ronsse et al., (2011) had participants learn to perform a 90 degree out-of-phase bimanual inward-outward wrist movement. One group received visual feedback in the form of a two-dimensional figure that showed their movement traces. A second group received auditory feedback instead in the form of sounds that marked the timing of particular points in their trajectories. Both groups improved their performance over time. The auditory group's retention test performance was unaffected even when auditory feedback was removed, but the visual group was severely impaired without visual feedback. This suggested that the visual group had become dependent on the feedback whereas the auditory group did not. This differential reliance on feedback is supported by the neuroimaging observation of remnants of visual activity even in the no-feedback retention test. The guidance hypothesis (Salmoni et al., 1984) holds that participants become

dependent on feedback and therefore performance returns to baseline once the feedback is removed. Neural correlates of motor learning under sensory feedback were further investigated using graph-theoretic network analyses (Heitger et al., 2012). In this study, a variety of network connectivity and integration metrics based on participant fMRI data were shown to improve during training. However, participants in the visual feedback condition showed a tendency for greater network connectivity. Although still under development, such cortical network metrics may help to document neural correlates of dependence on feedback in the brain. Furthermore, notice that these studies did not use a no-feedback baseline condition, in which participants would have no visual or auditory feedback to guide their learning. Therefore, strictly speaking, it has not been shown that feedback in itself is beneficial. For this reason, we will address this question in the present thesis (**part II**).

## **6.2 Sensory stimulation in action**

In this section, we discuss the influence of movement-independent sensory information (sensory stimulation) on action.

### **6.2.a Sensorimotor corticospinal facilitation**

The motor system can be primed by incoming sensory events in brief time windows. In this section, we discuss such responses for the auditory and visual modality. These phenomena have been advanced as explanations for the benefit of sound in motor rehabilitation (Malcolm et al., 2009; Thaut et al., 2002, 2007).

#### **i. Audiospinal facilitation**

Auditory-motor facilitation has been proposed to modulate the Hoffman's reflex (H-reflex). The H-reflex is measured as an EMG response of a muscle after electrically stimulating the sensory fibres (in the muscle spindles) of its innervating nerves. The response usually occurs 30 to 40 msec after the onset of the brief electric stimulation burst. It is similar to other observable reflexes such as the knee stretch reflex. The difference is that in the way the H-reflex is evoked: the muscle spindle receptors are bypassed by

stimulating the sensory nerves directly. For a recent review, see (McNeil et al., 2013).

The reflex is facilitated from about 50 to 250 msec after onset of an auditory stimulus, which has been dubbed audiospinal facilitation (ASF) (Davis et al., 1982; Davis and Beaton, 1968; Paltsev and Elner, 1967). Cortico-cortico circuits have shown to be involved in this process. Liegeois-Chauvel et al. (1989) showed that cortical lesions reduced or removed the facilitation (in the soleus muscle) altogether when the conditioning auditory stimulus was presented contra-laterally to the affected hemisphere. In particular, Heschl's Gyrus lesions appear to affect the ASF, suggesting that the neural circuitry involved passes through cortical pathways.

What is the purpose of this motor response to sound? Most initial studies used loud sounds that would startle the participants and therefore interpreted the response as an evolutionarily adaptive response to sudden changes in the environment that might signal danger (Davis et al., 1982; Rossignol, 1975). Later studies used less loud sounds and suggested audio-spinal facilitation to play a role in synchronisation to sounds (Rossignol and Melvill-Jones, 1976).

It is interesting to note that these studies have not used a learning paradigm. The measured effects were present immediately and did not depend on stimulus-response associations trained in participants (for example musical tone-to-key mappings). Similar non-specific auditory modulation of motor system excitability was shown using MEPs (but only for the dominant hand) (Mock et al., 2011). Therefore, the processes that these experimental paradigms tap into show that some connectivity between auditory and motor areas is inherent and not learned.

## **ii. Visual-induced motor modulation**

Startle-like reactions are not restricted to the auditory modality. It has been shown that salient visual stimuli can modulate motor excitability. For example, the sudden appearance of a (task-irrelevant) ball reduced MEP amplitudes in a visuomotor task as soon as 70-80 msec after its appearance, but only when it appeared at a location close to the hand (Makin et al., 2009). The finding suggest that our brain is exquisitely equipped to respond to sudden, potentially dangerous objects, preparing to move our limbs away from them.

Further, it provides support for the idea that the human brain contains limb-centered rather than (or at least in addition to) world-centered coordinates for visual objects (Graziano, 1999; Weidler and Abrams, 2013). Indeed, action planning and action execution are shown to bring forth a remapping of peri-personal space (Brozzoli et al., 2009) in a way that depends on the kind of action that is initiated (Brozzoli et al., 2010).

## **6.2.b Sensation as a reference for movement**

There is a wealth of examples from our everyday lives where we use the sensory system to guide action. For example, anecdotally, my driving teacher taught me to steer a curve by simply instructing me to look where I wanted to go. “Your movements that make the car go there,” he said, “will follow your gaze.” Scientific evidence for a perceptual basis for motor processing is abundant. For example, stability of anti-phase coordination of a bimanual cyclic finger movement is shown to depend on what the movement looks like rather than what the movement physically is (Mechsner et al., 2001).

A second example is that moving one's eyes smoothly (rather than in “jumps” - called saccades) is nearly impossible except when an object in the visual scene is moving regularly and we follow it with our eyes. This same smooth movement is very difficult to reproduce by participants in the absence of such an object (Krauzlis, 2005). Again, we find that the movement is guided by an external sensory object. Similarly, in the auditory domain, participants benefit from an auditory model when learning to produce a sequence of taps of irregular duration (Shea et al., 2001).

### **i. Periodic movement: sensorimotor synchronisation**

Sensorimotor synchronisation (SMS) refers to human's ability to synchronise one's movements to a periodic sensory stimulus (Repp, 2005). In a typical paradigm, participants are invited to tap their finger on a table in time with a metronome.

Trivially, motor and auditory systems are involved in the SMS task. It is, after all, the nature of the task: to synchronise motor actions with an auditory stimulus. In the case of sensorimotor synchronisation (and similar tasks) the surprising phenomenon is not that the auditory and motor systems are involved, but rather the fact that it is possible for the two to be used in concert in a way that is effective given the task at hand.

However, sensation may provide a temporal reference even when the task is not explicitly asking participants to synchronise their movements with an external pacing sequence. The bimanual advantage is the observation that when participants are asked to tap with their two hands as regularly as possible, the variability in timing of the individual hands is reduced relative to when they tap with only one hand at a time (Helmuth and Ivry, 1996). The finding is surprising because in bimanual tapping the brain needs to generate two movements and in unimanual tapping only one. One explanation for the phenomenon is that the brain maintains two independent timekeepers for the two hands which are then averaged, reducing their variance. Another explanation is that sensory reafferent codes provide a temporal reference (Drewing et al., 2002; Drewing and Aschersleben, 2003; Drewing et al., 2004). The reasoning for this explanation is as follows. Firstly, when two hands are involved, the auditory, tactile or proprioceptive sensory reafference is doubled. Secondly, the sensory information is thought to provide a control signal (a temporal reference). It then follows that having a richer sensory reference signal in the case of bimanual tapping enhances temporal precision. As predicted by this account, removing tactile sensory information by having participants tap in a opening in the table rather than on the table surface increased timing variability (Drewing et al., 2002). Secondly, as predicted, adding feedback tones to the left hand decreased timing accuracy of the right hand (for a discussion, see Drewing et al., 2004).

## ii. Patient populations

We now turn to clinical populations. Parkinson's disease (PD) patients are shown to benefit from a auditory pacing sequence in a finger tapping task and show a marked decrease in performance once this metronome is removed (Freeman et al., 1993). Similarly, patients benefit from an either auditory or visual pacing signal in gait (del Olmo and Cudeiro, 2003, 2005) and perhaps a combination of auditory and visual pacing (Suteerawattananon et al., 2004). Other studies (Lim et al., 2005) warn caution in interpreting methodologically problematic studies and point to the question of the extent to which the improvements generalise to patients' *activities of daily living* (ADL) that they so direly need improvements in.

Why do PD patients show a marked improvement when supplied with a pacing

signal? The dominant view is that their internal timekeeping mechanisms necessary for motor control are damaged by neural degeneration due to the disease. The external timekeeper provides an alternative pacing that can take the role of the missing internal one. Some indirect evidence for this idea is discussed in **study VII.1** of this thesis.

## 6.3 Speech production

Speech is a type of action that is of particular interest in the present context, because the auditory signal is the explicit goal of the movement, rather than a side effect of it (as it is in ripping a paper). This section discusses how humans rely on auditory feedback for online speech production.

### 6.3.a Sensory co-activation

A first line of evidence comes from the observation that silent reading typically activates the auditory cortex, relative to fMRI scanner noise listening baseline (MacSweeney et al., 2000; Capek et al., 2008). Similarly, silent reading and lip-reading modulated the auditory cortex' response to pure tones (Kauramäki et al., 2010). These findings show that even when the sensory consequence of action is absent, the corresponding cortical activation nevertheless obtains. This interpretation ties in with the observation that auditory feedback, when it *is* present, is suppressed relative to external auditory events (section 5.1.b).

### 6.3.b Reliance on feedback

In post-lingually deaf individuals, articulation quality is shown to decline slowly, as revealed by acoustic analyses. This indicates that auditory feedback is crucial in long-term maintenance of accurate speech control (Waldstein, 1990). By contrast, short-term effects of absence of auditory feedback are mild and scarcely documented. Humans are exquisitely sensitive to manipulations in the auditory feedback of their speech. For example, from our everyday lives we know how disrupting it is to speak on a phone or internet voice messaging when we hear our own voice with a delay (Hickok et al., 2011). Experimental

investigation confirms that speech fluency is severely affected by auditory feedback delay (Stuart et al., 2002; Yates, 1963) as well as by shifts in the spectrum of the produced sound (Stuart et al., 2008). When vowel formants are changed, participants respond by compensating for them (Cai et al., 2011; Shiller et al., 2009), and such compensation is greater across phoneme boundaries (Niziolek and Guenther, 2013). Children show similar compensation for feedback distortion but to a lesser extent, suggesting that auditory-motor connections may arise only later in life (Shiller et al., 2010).

Adaption to altered auditory feedback is less strong for patients with conduction aphasia. In this rare form of aphasia, patients have relatively intact speech comprehension and production capacities, but a deficit in repeating speech (Buchsbaum et al., 2011). Patients with conduction aphasia exhibit a decreased sensitivity to the delayed auditory feedback that is so disruptive for normal controls (Boller et al., 1978). Sometimes, patients even show an improved performance (Boller and Marcie, 1978). The paradox here is that healthy individuals show a conduction aphasia-like performance under delayed feedback, whereas conduction aphasics perform more normally. Similarly, stutterers are shown to benefit from delayed auditory feedback (Martin and Haroldson, 1979). It has subsequently been suggested that stuttering may be due to noisy feedback integration (Max et al., 2004); delayed feedback prevents this integration from affecting on-line speech production, thereby preventing the faulty feedback circuit's detrimental effect on overt speech (Hickok et al., 2011).

For a more detailed discussion of perceptual processes in speech production based on state feedback control architecture, see Hickok (2012).

## 6.4 Music production

Production of music shows parallels to speech: both are intricate, time-critical, communicative gestures aimed at the production of an auditory goal (Jäncke, 2012). Studying music production allows us to gain insight into this skill as it is being learned and gives the opportunity to compare experts with non-experts. In this section, we will focus on the influence of the perceptual system in music motor production.

#### 6.4.a Sensory co-activation during silent performance

Professional musicians report that when they play on a mute keyboard or when they are not able to hear their own instrument because of other loud instruments in their orchestra, they nevertheless are able to form a precise auditory image of what they sound like. Lotze et al. (2003) investigated violinists who imagined or actually produced the fingering movements of the beginning of a Mozart violin concerto. In both cases the movements were not performed on the violin and therefore did not yield their usual acoustic effects. Both professional and amateur violinists exhibited auditory cortex activation during the actual production of the fingering movements. In the professional violinists this activation was furthermore enhanced relative to amateurs, suggesting perhaps a greater auditory-motor coupling strength in these experts. Similarly, Meister et al. (2004) found bilateral premotor, precuneus and medial Brodmann Area 40 activation when music students imagined or overtly performed a familiar musical piece on a mute piano. As expected, actual performance additionally activated primary motor cortex and the bilateral posterior parietal cortex. These findings together with those from passive music perception (section 5.4) have led researchers to propose a common neural substrate involved in music perception and production (Bangert et al., 2006). However, this set of neural structures did not include auditory areas since they were not activated during mute music production. Professional musicians, however, revealed auditory cortex activation in a finger tapping task that was (deliberately) quite different from piano playing (Pau et al., 2013). Gebel et al. (2013) found that trumpet players without any sound emission yielded activation in auditory areas (Heschl's gyrus). What these studies suggest is that silent music production gives rise to mental auditory imagery of sorts (Zatorre, 2007; Zatorre et al., 1996), yielding cortical activations similar to those involved in the actual overt production.

In conclusion, there is evidence for auditory co-activation during silent music production. However, the fact that auditory activation is sometimes not found, or, paradoxically, *is* found in tasks very different from music playing, shows that further investigation is required to understand the nature of auditory co-activation. For example, studies could use motor-induced suppression-like paradigms (see section 5.1.b) to investigate auditory sensitivity in the temporal vicinity of silent keystrokes in professional musicians. If the auditory cortex activates as a result of a silent keystroke, this activity

should interfere with tone detection if the note is incongruent to the keystroke (e.g., a C note sounding when the participant silently played an E) or facilitate it when congruent (e.g., a D note sounding when the participant silently played a D). To our knowledge, such studies have not been performed.

#### **6.4.b Reliance on feedback**

Auditory feedback plays an important role in controlling on-line music performance (Chen et al., 2013a; Kajihara et al., 2013). Depriving pianists of auditory feedback influenced performance only marginally (Repp, 1999) or not at all (Gates et al., 1974; Finney, 1997; Finney and Palmer, 2003). However, tones being emitted with a (constant) delay severely deteriorates performance in a way that depends on the amount of delay (Gates and Bradshaw, 1974; Pfordresher, 2003; Pfordresher and Palmer, 2002), essentially mirroring observations in speech production discussed previously. Participants in non-musical tasks are shown to recalibrate the perceptual synchronicity of sensory effects when they are consistently presented later than the action that causes them (Keetels and Vroomen, 2012). It remains unclear whether musicians would similarly be able to adjust to delays in the long term.

A further common finding in finger tapping (in individuals with little or no musical experience) is that production is speeded up (for small delay in auditory feedback) or slowed down (for a longer delay) (Pfordresher and Dalla Bella, 2011). The latter study furthermore showed that the effect of the delay depended essentially on the position and velocity of the movement at the time of the (delayed) auditory feedback.

Taken together, the studies provide a convincing picture that auditory feedback does play a modulatory role in motor processes although they are probably not crucial for being able to perform the task.

#### **6.4.c Sonification feedback in musical novices**

Auditory feedback has also been generated for movements that originally do not produce sound or that originally produce a different sound. This artificial, augmented form of movement-to-sound mapping is called sonification. Strictly speaking, a musical

instrument is one form of sonification and therefore sonification studies can be viewed of ways of creating alternative musical instruments. The contribution of sonification to sensorimotor integration is that it becomes possible to study subtle manipulations in feedback and how they affect movement.

How are movements mapped to sound? There is an unlimited variety of such mappings and possibly experimental results depend on whether the sounds are aesthetically pleasing (Bergstrom et al., 2013), informative relative to movement and whether the mapping can be easily learnt. Sonification of brain signals has been developed to a much greater degree and may in the future provide inspiration for movement sonification techniques (for a survey, see Wu et al., 2009).

Passive observation of sonified swimming movements (of different individuals) was shown to activate a fronto-cortical network as well as the basal ganglia (BG) (Schmitz et al., 2013) suggesting that motor circuits may be readily activated. The authors also showed improved judgement of velocity profiles, but it remains unclear to what extent such passive perceptual improvement would generalise to benefits in active motor performance. Presently, the brain responses to active sonified movement are unknown.

It has been suggested that sonification may be beneficial for motor learning in populations that suffer from impairments in processing visual (Tan et al., 2013) or proprioceptive forms of feedback (such as stroke patients) or that need a high level of control (such as athletes). Alternatively, it is proposed that providing an additional stream of feedback may simply enhance sensitivity by enabling the brain to perform (Bayesian) integration of multiple (congruent) perceptual streams (Calvert et al., 2000; Shams and Seitz, 2008).

Healthy individuals are able to incorporate auditory feedback in the form of modulated pink noise in a task that required them to adapt their arm movements to disruptions in a force field (Oscari et al., 2012). The response to auditory feedback was largely comparable to that for visual feedback. In sports, sonification has been employed as a method to improve performance in professional rowers (Schaffert et al., 2011). The study compared rowing in normal conditions and when provided with an accelerometer-based sonification that was meant to make the different phases in the rowing movement more

salient. The boat went faster with this sonification than without. A participant remarked “[t]he sound pointed to things of which we have not been aware before and we became more conscious of the movement” (p.399). However, many questions remain open. For example, we wonder whether the mechanism that allowed the rowers to become more efficient was that the sound provided a common basis for synchronisation.

Although controlled investigation in this area remains sparse, the results above are promising and have propelled researchers to use sonification in clinical contexts finding encouraging results (Gonzalez et al., 2012; Inmann and Haugland, 2004; Janata and Edwards, 2013). The field is a rich ground for investigation into how the brain processes a sonification signal for the benefit of movement. For example, studies could investigate whether a sonification signal is integrated with motor information in the same Bayesian optimal way as visual information is (Kording and Wolpert, 2004).

## **I.7. Synthesis**

---

In this chapter, we have reviewed evidence for sensorimotor integration across a wide variety of modalities, placing an emphasis on audition. The thread of the discussion has been to investigate (a) motor involvement in perceptual tasks, and (b) involvement of the perceptual system in motor tasks. In both cases, the guiding questions have been what the role of the sensorimotor interaction is: what contribution does the one modality make to the performance of the other?

A wealth of experimental evidence points at motor involvement during passive perception. It is surprising that motor processes would at all be needed in a perceptual system that is supposed to simply objectively reflect the state of affairs in the world. Some theories have claimed that any thought is movement. Other theories proposed that motor activity is a process of internal simulation of the outside world, enabling the organism to understand, predict, interact or participate effectively. Yet other interpretations suggest that the motor activity might be an epiphenomenon, not serving any goal but resulting from repeatedly associating sensory events with motor events in a sort of Hebbian learning.

The other aspect of sensorimotor integration pertains to how the sensory system is involved in motor processes. It is agreed that perceptual information is used, both during motor skill acquisition and in maintaining of accuracy. In this way, perceptual information is used as guidance for a motor system that otherwise remains blind. A motor system not interacting with a perceptual module would not be able to do a variety of tasks (such as pressing a button) and those few tasks that it could do would have to be hard-wired into it and therefore clumsily rigid. Motor actions are internally simulated to yield predictions of the corresponding sensory consequences (possibly through internal models). These predictions are then used to (I) attenuate the incoming sensory information and (II) make temporal predictions of the occurrence of the movement that can subsequently be used for motor control and learning. This processing explains co-activation of sensory areas during (silent) motor activity as well as motor-induced suppression.

The two sides of sensorimotor integration have previously not been genuinely integrated. Theorists investigating sensory involvement in action have at best made sparse links with those investigating motor involvement in perception; and vice versa. If the underlying processes in these two sides are the same (which remains to be shown) then one expects that correlations between the critical empirical measures could be found.

A number of possible connections or dissociations between the various phenomena referred to by sensorimotor integration could be investigated experimentally. These could help to clarify the various ways in which sensory and motor areas interact. It is hoped that this chapter may provide a modest contribution by laying out a systematic conceptual map pointing to these connections or dissociations.

## **I.8. Outline of the thesis**

---

In this chapter, we have provided a review of sensorimotor integration. We show that it is a very broad term that is applied to many phenomena that are probably only loosely related. In particular, the literature is scattered and hardly interacts – or integrates. What needs to be known is whether the two sides of integration (motor-in-sensory and sensory-in-motor) are at all related or whether they are separate phenomena. This thesis

addresses this important discrepancy. We attempt to bring a key issue – that of the criticality of motor involvement in perception – from the motor-in-perception aspect into the perception-in-motor aspect. Concretely, we ask to what extent sensory involvement is crucial in motor learning. This question will be tackled by investigating a specific motor task, namely that of learning to time one's movements regularly.

The hypothesis of this thesis is that the criticality of the sensory involvement in motor learning depends on the population. Stroke patients with damaged auditory-motor systems may respond differently to alterations in auditory feedback. Expert musicians, who have benefited from auditory feedback for many thousands of hours perhaps do not need it any more (but may show evidence from having used it in the learning process).

The thesis comprises three parts. In part **part II** we will ask whether auditory feedback can be used in motor sequence learning in a healthy population. Based on its results, we investigate in **part III** how auditory feedback is used in non-musician stroke patients with motor deficits who are re-learning motor skills. Finally, in **part IV** we ask whether auditory feedback is still critical in expert musical performers.

## **II. Learning in musical novices**

## II.1. Introduction

---

In this part of the thesis, we investigate the role of auditory feedback in the learning of motor regularity. Our main questions are, firstly, whether auditory feedback can be used for motor learning by healthy non-musicians, and secondly, we ask how the timing between movement and sound is a critical element of learning.

This chapter introduces background literature for the experimental studies presented in this part. In section 1.1 we will define regularity in the context of motor sequence production and in section 1.2 we introduce a dominant paradigm used to study sequence learning, namely Serial Reaction Time (SRT). Our study II.3 will present a variant of this paradigm. Finally, we present the aims and hypotheses for this part of the thesis.

### 1.1 Motor regularity

We first define what regularity means in the context of motor sequence production. Imagine a task in which a participant is asked to press some sequence of keys. For example, this could be the case when the person plays a musical melody, but also when she types on a computer keyboard or enters a number in her phone. If we lay out the keystrokes over time, we can distinguish several possible scenarios (represented by the rows of Illustration 5). A participant can tap quickly (top two rows) or slowly (bottom two rows). Independent of this tapping speed, the regularity with which these keystrokes are produced can vary independently (Shimoyama, Ninchoji, & Uemura, 1990). For example, the participant can be regular (first and fourth rows) or irregular (second and third rows).

The metric typically used to quantify the regularity of a temporal sequence is calculated as follows. One computes the intervals between subsequent events (in our case, keystrokes) and calculates their standard deviation. However, if the intervals become shorter, the standard deviation tends to become smaller as well. That is, one needs to scale the irregularity by the length of the intervals in order to find a regularity estimate that is

maximally independent of speed (Shimoyama et al., 1990). Typically, researchers have calculated the coefficient of variation (CV) which is obtained by dividing the standard deviation by the mean. In this way, variability can be expressed as a percentage (although it can exceed 100%). At this point, it is important to recognise the underlying assumption that variability (defined as the standard deviation) varies *linearly* with speed. It may well be that variability is a different (non-linear) function of speed, but for the present purposes we will use this simplification.

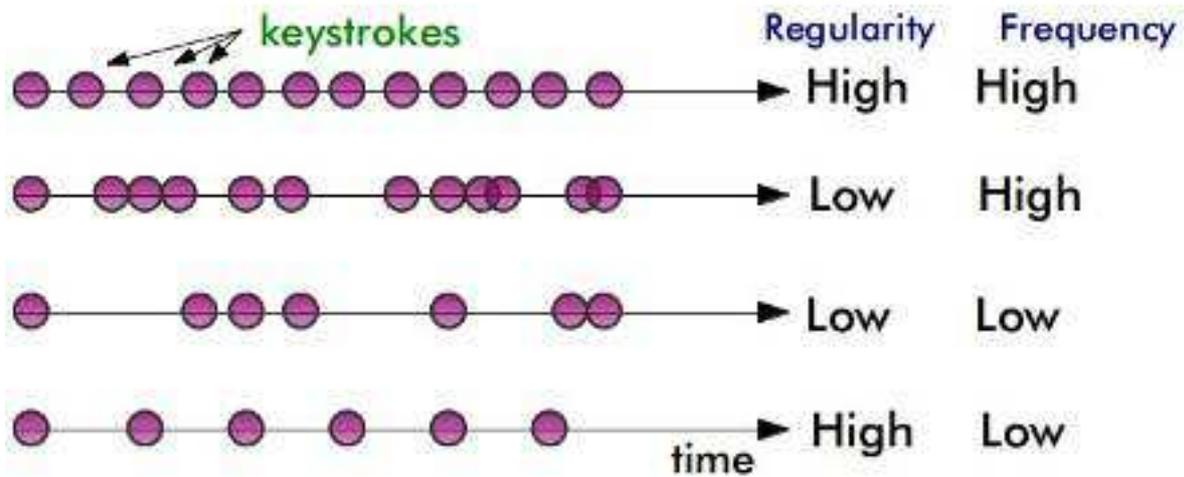


Illustration 5: Contrasting regularity and frequency (speed). Sequence production timing can, in principle, vary independently in speed and in regularity.

## 1.2 Serial Reaction Time

Motor sequence learning has been investigated in the framework of the Serial Reaction Time (SRT) (Nissen & Bullemer, 1987). Although our study II.3 will depart considerably from this framework, we will briefly discuss it here so that our manipulation can be better understood.

In a typical SRT study, participants need to respond to a sensory cue that can appear in several possible spatial locations. For example, four lights can be arranged in a row, out of which one will light up at any time. The participants respond as quickly as possible by

pressing the associated button, for example, one out of four buttons in a row. Human beings improve their performance as time passes as measured by a reduction in their response time (*sequence-unspecific learning* in Illustration 6). But the surprising thing is that if the order in which the cues appear is not random, and instead follows some pre-determined pattern over and over again, participants improve *more* than otherwise. This difference is measured by inserting a block in which the cues follow a different pattern than the pattern that was repeated on the previous blocks (*random test block* in Illustration 6). Suddenly, participants' reaction times increase dramatically, which is indicative that at least part of the learning has been specific to the sequence that they have practised. The finding is surprising because it shows that participants are able to pick up regularities in a sequence without being explicitly aware of this (Cleeremans, Destrebecqz, & Boyer, 1998). However, in our case, we will not deal with the question whether or not the sequence learning is implicit or explicit. In our experimental paradigm, we will always make explicit what the sequence is that participants are learning.

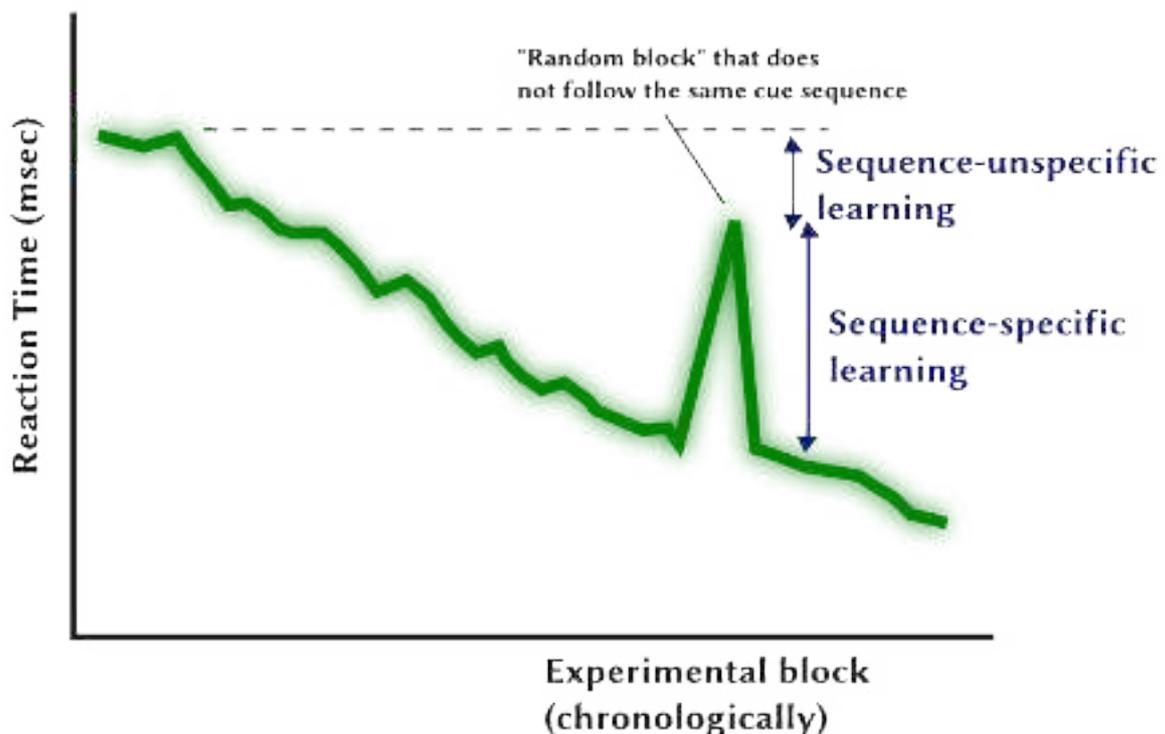


Illustration 6: Summary of findings in typical Serial Reaction Time studies.

### 1.3 Aims and Hypotheses

This part of the thesis is organised as follows. First, we will briefly introduce a test that we use to establish participants' thresholds for detecting timing differences between actions (in this case a keystroke) and their sensory results (in this case a tone). The test will be used in the second study in this part (study II.3) as well as elsewhere in this thesis (in study III.2). Next, we investigate whether the brain can use auditory feedback for motor learning, taking the learning of sequence production regularity as experimental paradigm. Finally, we ask what role is played by the timing between the action (the individual keystrokes) and the resulting effect (the sounds).

The aims of this part of the thesis are:

1. Introduce a simple test to measure humans' sensitivity to delays between their actions and sensory consequences (**study II.2**).
2. Establish whether auditory feedback can be used in learning motor regularity (**study II.3**). Our hypothesis is that the brain will be able to learn to tap a sequence more regularly when it receives sound feedback at each keystroke than when it does not.
3. Determine whether the timing of the feedback affects learning (**study II.3**). In particular, we hypothesise that if participants use the timing of auditory feedback, their learning should be affected by temporal disruptions in this feedback.

**II.2. Thresholds of auditory-motor coupling measured with a simple task in musicians and non-musicians: Was the sound simultaneous to the keystroke?**

---

## **Thresholds of auditory-motor coupling measured with a simple task in musicians and non-musicians: Was the sound simultaneous to the key press?**

van Vugt FT<sup>1,2\*</sup>, Tillmann B<sup>1</sup>

<sup>1</sup> Institute of Music Physiology and Musicians' Medicine, University of Music, Drama and Media, Hannover, Germany

<sup>2</sup> Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, CNRS-UMR 5292, INSERM U1028, University Lyon-1, Lyon, France

### **Abstract**

The human brain is able to predict the sensory effects of its actions. But how precise are these predictions? The present research proposes a tool to measure thresholds between a simple action (keystroke) and a simultaneously emitted sound. On each trial, participants were required to press a key. Upon each keystroke, a woodblock tick sound was presented over headphones. In some trials, the sound came simultaneous with the downward keystroke; at other times, it was delayed by a varying amount of time. Participants were asked to verbally report whether the sound came immediately or was delayed. Participants' delay detection thresholds (in msec) were measured with a staircase-like procedure known as Maximum Likelihood Procedure. We hypothesised musicians would have a lower threshold than non-musicians. Comparing pianists and brass players, we furthermore hypothesised that, as a result of a sharper attack of the timbre of their instrument, pianists might have lower thresholds than brass players. Our results show that non-musicians exhibited higher thresholds for delay detection ( $180 \pm 104$  msec) than the two groups of musicians ( $102 \pm 65$  msec), but there were no reliable differences between pianists and brass players. The thresholds correlated with as well as with sensorimotor synchronisation capacities and with a purely auditory threshold measure requiring the detection of a temporal irregularity in an otherwise isochronous sound sequence. These findings suggest that the brain has a relatively large window of integration within which an action and its resulting effect are judged as simultaneous. Furthermore, musical expertise may narrow this window down, potentially due to a more refined temporal prediction. This novel paradigm provides a simple test to estimate the temporal precision of auditory-motor

action-effect coupling that can readily be incorporated in studies investigating both healthy and patient populations.

## 2.1 Introduction

Many motor actions have sensory consequences. For example, we see our hands displace when we move them, and our steps make sounds. The human brain is able to predict the sensory effects of its actions (Blakemore, Rees, & Frith, 1998; Blakemore, Wolpert, & Frith, 1998; Eliades & Wang, 2003). These predictions are crucial for distinguishing between sensory information that is generated by oneself and sensory information coming from outside. In particular, self-produced sensory effects are suppressed in comparison with externally produced effects (Martikainen, Kaneko, & Hari, 2005).

The brain is able to predict not only *what* sensory event will follow its action, but also *when* it is supposed to occur. This is evident from the observation that self-produced sensory effects are no longer suppressed when they are delayed by several hundreds of milliseconds (Aliu, Houde, & Nagarajan, 2009). This temporal prediction is not fixed, but adaptive to the situation. For example, the point of subjective synchrony (PSS) between various sensory events can be recalibrated, even to the extent that the physical order of events can be inverted (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Kuling, van Eijk, Juola, & Kohlrausch, 2012; Tanaka, Asakawa, & Imai, 2011; S. Yamamoto, Miyazaki, Iwano, & Kitazawa, 2012). Lesions may affect subjective synchrony as shown by the intriguing case of a man who hears people speak before their lips move (Freeman et al., 2013). Synchrony can also be recalibrated between sensory and (active) motor events (Keetels & Vroomen, 2012; Rohde & Ernst, 2012; Sugano, Keetels, & Vroomen, 2010; K. Yamamoto & Kawabata, 2011).

But how precise are these predictions and the perceived synchrony between pairs of events (sensory-sensory; sensory-motor)? Common experimental paradigms to measure this precision are asking participants to either judge whether two stimuli are simultaneous (simultaneity judgement task - SJ) or report the order of two stimuli (temporal order judgement - TOJ). For the temporal order judgement task (TOJ), precision is measured as

the just-noticeable difference (JND) between two potential orderings. Humans can distinguish two auditory clicks presented to the same ear when they are separated by 2 msec, but at least 60 msec are needed to distinguish them binaurally (Exner, 1875). Typical thresholds for TOJ between two auditory stimuli are inter-stimulus-intervals (ISIs) of 20 to 60 msec, probably depending on the stimulus type (Ben-Artzi, Fostick, & Babkoff, 2005; Fostick & Babkoff, 2013; Szymaszek, Szelag, & Sliwowska, 2006). Thresholds for TOJ between auditory (tone) and visual (flash) stimuli are typically anywhere between 25 and 50 msec (Hirsh & Sherrick Jr, 1961; Zampini, Shore, & Spence, 2003). Auditory-haptic thresholds usually have JNDs of 100 msec, and haptic-haptic thresholds have JNDs of around 50 msec (Frissen, Ziat, Champion, Hayward, & Guastavino, 2012). Although the SJ and TOJ tasks often give different results, thresholds for the SJ task tend to be smaller than those for the TOJ task (García-Pérez & Alcalá-Quintana, 2012; Weiss & Scharlau, 2011). This led to the dominant view that they probably reflect different underlying processes (Vatakis, Navarra, Soto-Faraco, & Spence, 2008). Training plays a role in shaping sensitivities, as is shown by video game players having smaller thresholds for audio-visual simultaneity judgements than non-players (Donohue, Woldorff, & Mitroff, 2010).

It remains unclear how sensitive participants are to the synchrony between events that they actively produce (such as keystrokes) and their sensory consequences (such as tones). Previously, this question has been studied by investigating the effects of altered sensory feedback to a produced action. For instance, musicians' timing performance was measured when they played on a piano that emitted the played sounds with a delay. Large delays (such as 200 msec) are noticeable and disrupt the fluidity of performance (Gates, Bradshaw, & Nettleton, 1974; Pfordresher, 2003; Pfordresher & Palmer, 2002). Speakers' fluency is similarly affected when auditory feedback is delayed (Stuart, Kalinowski, Rastatter, & Lynch, 2002; Yates, 1963). In order to be able to assess quantitatively whether disruptions in auditory feedback are noticeable and to investigate the effect of training and expertise, there is a need for an experimental paradigm that can establish the thresholds for action-effect synchrony judgements.

The present research proposes a new tool to measure thresholds between a simple action and an emitted sound. In this task, participants are asked to press a key. Either immediately or after a predetermined duration has elapsed, they are presented a sound

through their headphones. Our aim was to measure the thresholds for detecting a delay between the keystroke and the sound, and to investigate the effect of expertise. In addition, our aim was to establish how this action-effect synchrony sensitivity relates to other auditory and auditory-motor capacities. To this end, our participants also performed, firstly, an auditory temporal deviant detection task, and secondly, a sensorimotor synchronisation task. That is, we measured how well they could synchronise their movements to an external stimulus (Repp, 2005; Repp & Su, 2013). For this, we used a variation on the synchronisation-continuation tapping paradigm (Tillmann, Stevens, & Keller, 2011; Wing & Kristofferson, 1973). All tasks were performed by non-musicians and by pianist and brass player musicians. It has been previously reported that musicians outperform non-musicians with increased auditory discrimination (Ehrlé & Samson, 2005; Yee, Holleran, & Jones, 1994), and by tapping closer to the beat and more precisely (Aschersleben, 2002; Repp, 2004). We further hypothesised that the relation between finger movements and sounds for the musicians' main instrument might influence the thresholds too: when pianists strike a key the sound is instantaneous, whereas brass players' sound onset is determined by their respiration. Also, the piano sound has a sharper onset than the brass sound. As a result, we expected that pianists would have lower thresholds than brass players. We had also considered singers as an alternative to the brass players, but found that they tend to have a large amount of piano training as secondary instrument, which would confound the comparison with pianists.

## **2.2 Materials and Methods**

### **2.2.a Participants**

We recruited two groups of musician participants from the student pool at the Hanover Music University and young professionals. We furthermore recruited non-musicians in the same age range. Table 1 lists biographical and questionnaire data of each group. Participants self-reported no hearing impairment or neurological disorder, were aged between 18 and 40 years and right-handed. The musician participants were recruited in two groups: one group whose major study topic was piano (or who were professional

pianists) and another group who majored in a brass instrument. Inclusion criterion for the non-musician group were having received less than 1 year of musical training (apart from obligatory courses in primary or secondary schools).

Among the brass players, 13 had received piano instruction in the form of obligatory courses at the conservatory or in their childhood. For the entire brass group, the lifetime accumulated piano practice time was 1.1 (SD 1.8) thousand hours over an average total of 6.9 (SD 6.1) years.

Participants filled out a questionnaire with basic information such as age, handedness (according to the Edinburgh Handedness Inventory), and instrumental practice prior to their participation (Table 1). We found that basic biographical parameters did not differ except for computer keyboard use (Kruskal-Wallis  $\chi^2(2)=7.84$ ,  $p=.019$  uncorr.). This effect indicated that the musician group reported more daily computer keyboard use than the pianists [Mann-Whitney  $U=98.0$ ,  $Z=2.55$ ,  $p=0.01$ ,  $r=0.07$ ] or brass players [Mann-Whitney  $U=96.5$ ,  $Z=2.20$ ,  $p=0.03$ ,  $r=0.06$ ]. The two musician groups did not differ in their use of computer keyboards [Mann-Whitney  $U=161.0$ ,  $Z=0.61$ ,  $p=0.54$ ,  $r=0.02$ ].

	Pianists	Brass	Nonmusicians	Comparison between groups
N	20	18	18	
Gender (female/male)	10/10	7/11	8/10	$\chi^2(2)=.47, p=.79$
Age (years)	26.1 (5.7)	24.9 (3.5)	26.2 (4.7)	Kruskal-Wallis $\chi^2(2)=1.07, p=.59$
Handedness (Handedness Quotient in %)	73.4 (19.9)	75.3 (16.5)	78.1 (20.0)	Kruskal-Wallis $\chi^2(2)=1.10, p=.58$
Capable of blind typing (number of participants in each of the following categories: 10 fingers/less than 10 fingers/none)	2/14/4	7/10/1	5/10/3	Kruskal-Wallis $\chi^2(2)=4.67, p=.10$
Video game use in hours per week (number of participants in each of the following categories: none/ <1h / 1-7h / >7h)	16/3/1/0	10/7/1/0	13/1/3/1	Kruskal-Wallis $\chi^2(2)=2.09, p=.35$
Use of computer keyboards in hours per day (number of participants in each of the following categories: <1h / 1-2h / >2h)	10/8/2	7/9/2	5/2/11	Kruskal-Wallis $\chi^2(2)=7.84, p=.019^*$
Capacity in using computer keyboards (self-rated 1-10)	6.3 (1.6)	6.6 (1.9)	6.7 (2.1)	Kruskal-Wallis $\chi^2(2)=0.87, p=.65$
Use of text messaging on cell phone in hours per week (number of participants in each of the following categories: none/ <1h / 1-7h / >7h)	7/9/4	4/12/2	10/4/4	Kruskal-Wallis $\chi^2(2)=1.42, p=.49$
Capacity in using text messaging (self-rated 1-10)	7.3 (2.0)	6.8 (1.9)	6.2 (2.3)	Kruskal-Wallis $\chi^2(2)=2.41, p=.30$
Age of onset of musical training (years)	6.65 (2.2)	9.78 (3.1)	NA	$t(30.5)=-3.56, p=.001^{**}$
Accumulated practice time on principal instrument (x10,000 hours)	22.6 (10.5)	13.1 (8.1)	NA	$t(35.3)=3.15, p=.003^{**}$
Years of musical practice	19.5 (5.6)	15.1 (3.6)	NA	$t(32.6)=2.90, p=.007^{**}$
Current daily practice time (hours)	3.7 (2.2)	3.3 (1.8)	NA	$t(35.6)=0.68, p=.50$
Absolute hearing (yes/no; self-reported)	7/13	0/19	NA	Fisher Exact Test $p=.009^{**}$

Table 1: Basic information about the three groups of participants. Data is reported as mean (SD) unless otherwise specified. Uncorrected significance is indicated: \* $p<.05$ , \*\* $p<.01$ , \*\*\* $p<.001$

## 2.2.b Materials

### i. Keystroke-sound delay detection task

We used a USB keypad (Hama Slimline Keypad SK110) that interfaced through HDI protocols with a python script detecting the keystroke onset and playing a woodblock wave sound (duration: 63 msec) through headphones (Shure SRH440). The woodblock sound was chosen because of its relatively sharp sound onset and nevertheless being pleasant to hear.

### ii. Anisochrony detection

We used a python-pygame graphical user interface that presented the sounds and the instructions. Instructions were given orally as well. The five-tone sequences were generated as follows. The base sequence consisted of five isochronous sine wave tones of 100ms presented with an inter-onset-interval (IOI) of 350ms. In some trials, the fourth tone was delayed by a certain amount but the fifth tone was always on time (Ehrlé & Samson, 2005; Hyde & Peretz, 2004). That is, when the tone was delayed by an amount  $d$ , the third interval was longer by  $d$  msec and the fourth interval was shorter by  $d$  msec.

### iii. Synchronisation-Continuation Tapping

The synchronisation stimulus was generated offline as follows and saved to a wave file. First, we created 4 finger snap sounds with an inter-onset-interval of 300 msec. Then followed 30 instances of the woodblock sound (the same as used during the learning part of the experiment) with an inter-onset-interval (IOI) of 600 msec. This was followed by a silence of  $30 \times 600$  msec (the equivalent of 30 more taps). Finally, a high-pitched gong sound was used to signal the end of the trial. The sounds were played using a custom developed python experimental script, which also communicated through a HID-USB interface with the button box to register the responses.

Participants' finger taps were recorded using a custom tapping surface containing a (piezo-based) contact sensor that communicated with the computer through a serial interface and was captured in a python program that also presented the stimuli (using pyAudio).

## 2.2.c Procedure

### i. Keystroke-sound delay detection task

In the delay detection task, we measured participants' sensitivity to delays between motor (keystroke) and auditory (tone) events. That is, we established from which delay onwards participants noticed that the tone came after the keystroke instead of immediately. At each trial, the participant pressed the “zero” key on the keypad at a time of her/his choosing and heard a tone. This tone was either played at the same time of the keystroke or delayed by a number of milliseconds. The participants responded verbally whether or not they had the feeling that the tone was delayed. Their responses were entered in the computer by the experimenter. Crucially, participants were instructed to leave their finger on the key (instead of lifting it prior to the keystroke) so as to reduce the tactile timing information. Furthermore, they were required to keep their eyes closed during the keystroke.

We used the Maximum Likelihood Procedure (MLP) algorithm to detect the threshold for the detection of the asynchrony between movement (keypress) and the tone. The algorithm is designed to adaptively select the stimulus level (tone delay) on each trial so as to converge to the participants' threshold. For each block, the algorithm outputs an estimate for the participant's threshold.

The MLP algorithm briefly works as follows. A set of candidate psychometric curves are maintained in parallel and for each, the likelihood of the set of the participants' responses is calculated. The psychometric curve that makes the participant's responses maximally likely is used to determine the stimulus level (the delay between the keystroke and the sound) on the next trial. We used 600 candidate psychometric curves with midpoints linearly spread between 0 and 600ms delay and crossed by the five false alarm rates (0%, 10%, 20%, 30%, 40%). Hence, a total of 3000 candidate psychometric curves were used.

Participants first performed 4 trials (2 with no delay and 2 with a delay of 600ms) to make clear the difference between when the sound came immediately and when it was delayed. The participant received accuracy feedback about her answers during these practice trials. Next, they performed a block of 10 trials, starting at 600ms delay but then

using MLP to determine the stimulus levels of the following trials. If the procedure was clear, we continued with 3 experimental blocks of 36 trials. Each experimental block consisted of 36 trials containing 6 catch trials. Catch trials are trials on which the delay was always 0 msec (regardless of the delay that was suggested by the MLP algorithm). The function of catch trials is to prevent participants from always responding “delayed” (which would cause the MLP algorithm converge to a zero threshold). Catch trials were inserted randomly as follows: the first 12 trials contained 2 catch trials and the next 24 trials contained 4 catch trials.

The maximum likelihood procedure was implemented in python. We made our source code freely available online on <https://github.com/florisvanvugt/PythonMLP>.

## ii. **Anisochrony detection**

Participants were seated comfortably and on each trial heard a sequence of five tones (see materials). Participants' task was to respond whether the five tone sequence was regular or not by pressing one of two response keys on the laptop keyboard. Stimuli (see materials) were presented through headphones set to a comfortable sound level that was kept constant across all participants. The participant's threshold was established adaptively using the on-line maximum-likelihood procedure (MLP). The basic procedure was the same as for the delay detection task, but here the set of candidate psychometric curves was as follows. We defined 200 logistic psychophysical curves whose midpoints were linearly spread over the 0 to 200ms delay range (0% to 57% of the tone IOI) and these were crossed with the five false alarm rates (0,10,20,30,40%). Again, each experimental block consisted of, first, 12 trials containing 2 catch trials, and then 24 trials containing 4 catch trials.

After the participant read the instructions, the interface presented the four example stimuli (two regular, two irregular). For these trials participant received accuracy feedback. The first trial of the next block of 10 trials was set to a keystroke-sound 200 msec delay and then used the adaptive procedure (MLP) to determine the stimulus level on the next trials. During this second training block, no accuracy feedback was provided. Finally, when the instructions were clear, participants continued with three experimental blocks. In between blocks, participants took a brief break of several minutes.

### iii. Synchronisation-Continuation Tapping

In each trial, participants tapped with their index finger on a flat surface along with the synchronisation stimulus after the four finger snap sounds (see materials). When the woodblock sounds stopped, participants were instructed to continue tapping at the same speed and regularity until the high-pitched sound signaled the end of the trial.

### iv. Data analyses

The threshold tasks were analysed as follows. First, we discarded blocks that contained more than 30% incorrect catch trial responses (in which the delay or deviation was 0 msec). Secondly, we discarded blocks in which the threshold estimate had not properly converged towards the end of the block. This was tested by fitting a regression line to the last 10 trials in the block, and discarding those blocks in which the slope of this line exceeded 2 msec/trial (for the delay detection task) or 1.18 msec/trial (for the anisochrony task). These slope cut-off points were chosen so as to firstly confirm visual inspection of blocks that had not properly converged and secondly to be roughly the same proportion of the average final threshold in the anisochrony and delay detection task. Thirdly, we computed the average threshold estimate for the remaining blocks for each participant.

Synchronisation tapping performance was analysed using linear and circular statistics. In the linear analysis, we calculated the time between each tap and its corresponding metronome click (in msec). For each block, we averaged these to yield the mean relative asynchrony (in msec) and calculated the standard deviation to yield the SD relative asynchrony (in msec). The mean relative asynchrony is a measure for how close participants tapped to the beat and the SD relative asynchrony is a measure of tapping precision (time-lock). In the circular analysis (Fisher, 1995), the timing of each tap was converted into a phase (between 0 and  $2\pi$ ) relative to the metronome onset. Based on these, we calculated the synchronisation vector, which is the average of all vectors with length 1 and the phase angle for that tap. The length of this vector (between 0 and 1) is a measure for the time-lock between the tap and the sound. In order to average these vectors within groups, we used Fisher's r-to-z transformation and obtained the corresponding z-scores.

For the continuation phase (when the metronome had stopped), we calculated the intervals between taps (inter-tap-interval, ITI, in msec) and its standard deviation (SD ITI in msec). We then de-trended the continuation taps by fitting a regression line to the ITIs over time, reporting the slope of this line and taking the residual variability from this line. In this way, we compensated for the fact that participants tend to speed up or slow down (Drewing, Stenneken, Cole, Prinz, & Aschersleben, 2004; Helmuth & Ivry, 1996; Keele, Pokorny, Corcos, & Ivry, 1985). The slope of this line fit indicated the tempo drift.

In order to compare performance of the three groups, we performed between-participants ANOVAs. We tested for homogeneity of variance using Levene's Test, and report where it was significant. We report generalised effect sizes (Bakeman, 2005). Follow-up comparisons were calculated using Tukey's HSD method.

## 2.3 Results

### 2.3.a Delay Detection

We discarded 17.0% of all blocks because of catch trial errors, and a further 2.3% because of lack of threshold convergence. Four participants (2 pianists, 2 brass players) had no remaining blocks and were eliminated from further analyses. For the other participants, we calculated the average of the thresholds on the basis of the remaining 2.6 (SD 0.7) blocks.

The distribution of thresholds of all participants in all groups combined was significantly non-normal [Shapiro-Wilk normality test  $W=.86$ ,  $p=.00003$ ], and therefore we continued statistical analyses with log-transformed thresholds. These did not violate normality assumptions [Shapiro-Wilk  $W=.98$ ,  $p=.71$ ]. The main effect of group (pianist, brass, nonmusician) on delay detection threshold was significant [ $F(2,49)=6.40$ ,  $p=.003$ ,  $\eta_G^2=.21$ ]. Post-hoc Tukey HSD contrasts indicated that the nonmusicians' threshold was higher than those of the pianists [ $p=.01$ ] and than those of the brass players [ $p=.006$ ]. The brass players and pianists' thresholds were not significantly different [ $p=.93$ ] (Figure 1 left). Among the brass players, we found that those who played piano as their second instrument had a lower delay detection threshold ( $M=83.0$ ,  $SD=42.5$ ) than those who did not ( $M=116.2$ ,

SD=70.0). However, this difference was not significant [ $t(7.3)=1.09$ ,  $p=.31$ ].

### 2.3.b Anisochrony

We discarded 11.1% of all blocks because of catch trial errors, but no further blocks were discarded because all had properly converged. Two participants (1 brass, 1 pianist) had no blocks remaining (based on the first criterion) and were eliminated from further analysis. For the other participants, we averaged the remaining 2.7 (SD=0.6) blocks into a single threshold value per participant.

The distribution of thresholds was significantly non-normal [Shapiro-Wilk normality test  $W=.92$ ,  $p=.0009$ ] and therefore we continued statistical analyses with log-transformed thresholds. These did not violate normality assumptions [Shapiro-Wilk  $W=.97$ ,  $p=.20$ ]. The main effect of group on anisochrony threshold was significant [ $F(2,51)=21.60$ ,  $p<.0001$ ,  $\eta_G^2=.46$ ]. Tukey HSD contrasts indicated that the nonmusicians' threshold was higher than those of the pianists [ $p<.001$ ] and than those of the brass players [ $p<.0001$ ]. The brass players and pianists' thresholds were not significantly different [ $p=.52$ ] (Figure 1 right). In the pianist group, there was one outlier who was further than 3 SD away below the mean for that group, but removing this participant did not affect any of the results.

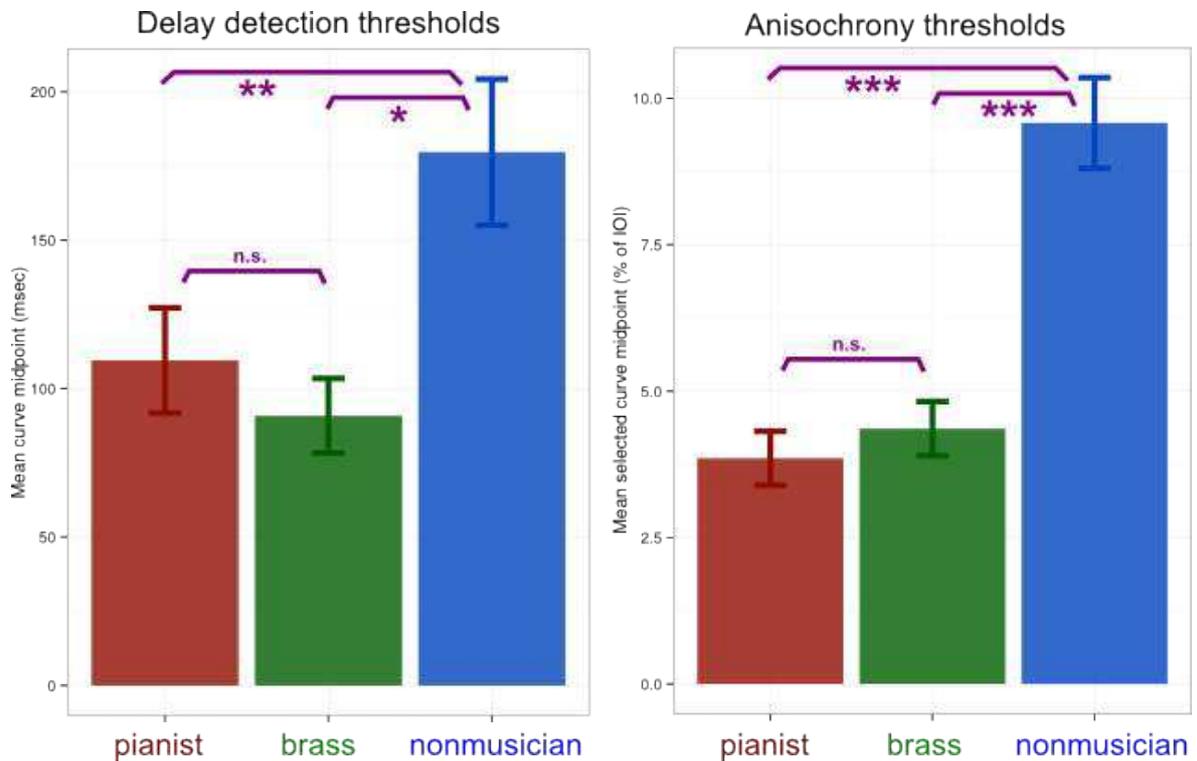


Figure 1: Results of the auditory and motor tests. The figures indicate the average thresholds for each of the groups (error bars indicate the standard error of the mean). \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

### 2.3.c Synchronisation-Continuation Tapping

We report basic measures of synchronisation and tapping variability in Table 2. Tukey contrasts revealed that brass and pianists do not differ in any of the measures (all  $p > .73$ ) but contrasts between the non-musicians on the one hand and the and pianist and brass groups on the other yielded significant or marginally significant differences (all  $p < .08$ ) (Table 2).

	Pianists	Brass players	Nonmusicians	Between-group comparison
<i>Synchronisation phase</i>				
Mean relative asynchrony (msec)	7.5 (21.2)	5.7 (28.3)	-22.5 (52.3)	F(2,44)=3.35, p=.04*
SD relative asynchrony (msec)	19.9 (4.9)	19.3 (3.2)	37.4 (20.3)	F(2,44)=11.7, p<.0001
Synchronisation vector length (r-bar, z-transformed)	2.3 (0.2)	2.3 (0.2)	1.8 (0.4)	F(2,44)=20.85, p<.00001
<i>Continuation phase</i>				
Continuation ITI (msec) (without detrending)	604 (10)	605 (11)	596 (20)	F(2,44)=1.79, p=.18 <sup>+</sup>
Continuation SD ITI (msec) (without detrending)	17.9 (2.9)	19.6 (2.7)	31.4 (9.0)	F(2,44)=27.04, p<.000001
Continuation drift (msec/sec)	-0.3 (0.6)	-0.4 (0.8)	-0.9 (1.0)	F(2,44)=2.65, p=.08.
Continuation residual variability after detrending (CV %)	5.5 (0.9)	6.0 (0.8)	9.7 (2.9)	F(2,44)=25.7, p<.00001, $\eta^2=.54$

Table 2: Synchronisation and continuation tapping results for the three groups. Values are reported as mean (SD) unless otherwise specified. <sup>+</sup>For the continuation ITI, Levene's test for homogeneity is violated.

### 2.3.d Comparison between the tests

Participants' performance on the tests reported here were not independent. Combining the thresholds from the three groups, the delay detection threshold correlated positively with the anisochrony task [Pearson  $\rho(49)=.60$ ,  $p<.0001$ ,  $R_{adj}^2=.35$ ]. The delay detection threshold also correlated negatively with the synchronisation vector length [Pearson  $\rho(49)=-.53$ ,  $p<.0001$ ,  $R_{adj}^2=.27$ ].

To test whether these correlations differ statistically between the groups, and whether performance on the anisochrony and synchronisation tasks combined might explain more of the variance in delay detection than either of those two tasks alone, we

performed the following analyses. Participants who had at least one valid anisochrony block and at least one valid delay detection block remaining (after discarding) entered in this analysis. This was the case for 17 pianists, 16 brass players and 18 non-musicians. We ran a linear ANCOVA model with delay detection threshold as dependent variable, group (nonmusician, brass player or pianist) as categorical factor (between-participants) and anisochrony threshold and sensorimotor synchronisation accuracy (vector length  $\bar{r}$ ) as covariates.

The interaction between anisochrony threshold and group was not significant [ $F(2,48)=1.64, p=.21$ ], which indicated that the linear relationship between the anisochrony and delay detection thresholds were not different between the groups. The interaction between synchronisation accuracy and group was not significant either [ $F(2,48)=.91, p=.41$ ]. This means that the linear relationship between synchronisation accuracy and delay detection was not different between groups. The main effect of anisochrony threshold was significant [ $F(1,48)=5.56, p=.02$ ] as was the main effect of synchronisation accuracy [ $F(1,48)=8.73, p=.004$ ]. There was no main effect of group [ $F(2,48)=1.06, p=.35$ ]. These results were essentially the same when repeated without the participant with an outlier anisochrony threshold (Figure 2).

In sum, the anisochrony and synchronisation accuracy both significantly explained the variance in delay detection thresholds (Figure 2). Taken together, they explained more than either one factor alone. With these two predictors, the group (pianist, brass, nonmusician) factor did not explain additional variance, indicating that the musicianship effect on the delay detection threshold was explained by anisochrony and synchronisation task performance.

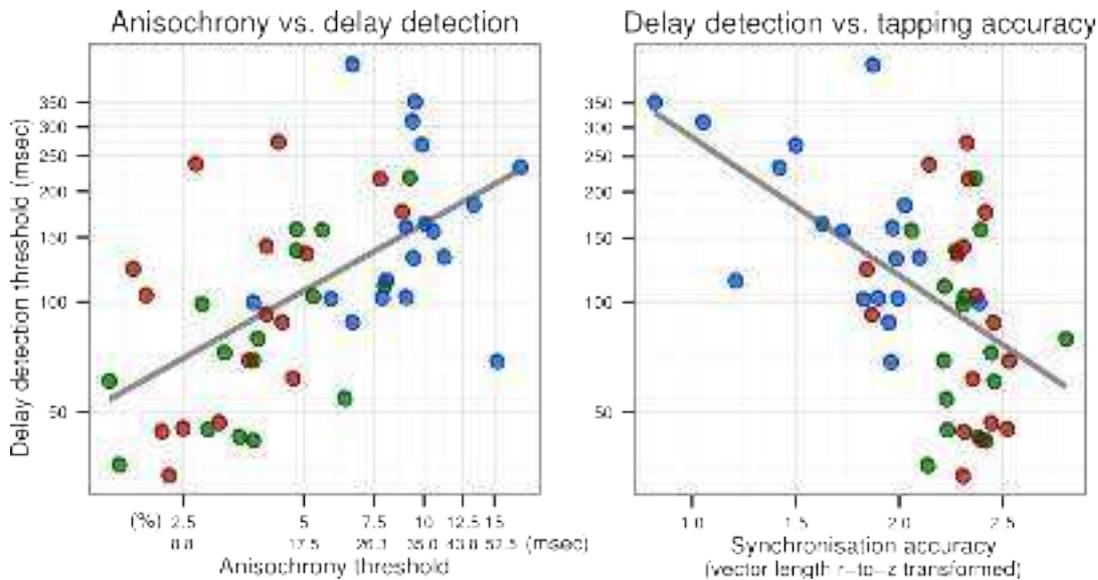


Figure 2: Correlations between the delay detection, anisochrony and tapping tests. The dot colour indicates the group: blue for non-musicians, red for pianists and green for brass players.

## 2.4 Conclusion

The human brain predicts sensory effects of its motor actions (Blakemore, Rees, et al., 1998; Friston, 2012). Not only does the brain predict *what* effect will follow, but also *when* it is expected to occur (Aliu et al., 2009). The present paper presents a simple test to measure the precision of this temporal prediction window. We applied this test to a non-musician population and two groups of musicians: brass players and pianists in order to investigate the effect of training. We furthermore asked how the sensitivity to auditory-motor delays builds on other auditory and auditory-motor tasks.

Our findings suggest that the brain has a relatively large window of integration ( $102 \pm 65$  msec for musicians, and  $180 \pm 104$  msec for nonmusicians?) within which an action and its resulting effect are judged as simultaneous. These delay detection thresholds are larger by almost an order of magnitude than thresholds for judging two auditory events as asynchronous, which are between 2 and 60 msec (Ben-Artzi et al., 2005; Exner, 1875; Fostick & Babkoff, 2013; Szymaszek et al., 2006). However, the present findings are in line with cross-modal sensory asynchrony judgements: visual and auditory events simultaneity

thresholds are usually around 150msec (Stevenson & Wallace, 2013).

Participants' capacity to judge simultaneity of movement and sound was explained as a combination of auditory temporal prediction precision (anisochny) and sensorimotor synchronisation accuracy. That is, the delay detection task appears to rely on cognitive capacities of basic auditory processing and auditory-motor coupling. Both of these capacities varied with musicianship, which did not additionally explain thresholds for audio-motor synchrony judgements. These results indicate that, first of all, sensitivity to auditory-motor delays can be trained. Musicians were more precise in temporal predicting the auditory effect of their movement, as evidenced by their lower threshold in the delay detection task. This finding is in line with the finding that musical training improves performance in a variety of tasks (Aschersleben, 2002; Ehrle & Samson, 2005; Repp, 2004; Yee et al., 1994) and also induces functional and structural brain changes (Gaser & Schlaug, 2003; Herholz & Zatorre, 2012). In addition, the finding is in line with previous studies showing that temporal order judgements (TOJ) improve with training (Alais & Cass, 2010; Powers, Hillock, & Wallace, 2009). Secondly, musicianship appears to improve delay detection thresholds indirectly. Musicianship did not significantly influence delay detection sensitivity when performance on purely auditory (anisochny) or auditory-motor tasks (sensorimotor synchronisation) was taken into account. This means that auditory-motor delay detection is not a capacity that is specifically improved by music. If this were so, we would have expected to find differences in correlations between the tests (delay detection, anisochny and sensorimotor synchronisation) between our groups. This was not the case. Instead, we propose that musical training improves sensorimotor synchronisation capacities as well as auditory temporal precision, both of which then lead to an improvement in delay detection threshold. Furthermore, the instrument that musicians played had no influence on delay detection sensitivity, or any of our other tasks. This suggests that the specifics of how an instrument responds to finger movements of the musician nor the acoustic features of the instrumental sound influence the capacity to detect delays between movement and sound.

Humans' sensitivity to delays between their articulator movements and the produced speech sound is typically around 60-70 msec (K. Yamamoto & Kawabata, 2011). These delays are below thresholds observed here, but closer to the thresholds we found for

musicians. Humans accumulate many hours of speech practice (many more than even professional musicians could accumulate on their instrument) and therefore one will expect to find lower delay detection thresholds for vocal actions. This finding squares with the idea that training an action, be it speaking or playing an instrument, improves the temporal prediction of its sensory consequences. However, the particular instrument that the musicians trained to play (piano or brass instruments) did not influence sensitivity, suggesting that perhaps, delay sensitivity is specific to the effector: the articulators in the case of speech and the hand in the case of piano playing and perhaps the mouth in the case of brass playing. Notice, however, that comparisons between music and speech are limited by the fact that there exist no control group with negligible speech experience.

In sum, the present findings suggest that the brain has a relatively large window of integration within which an action and its resulting effect are judged as simultaneous. Furthermore, musical expertise may narrow this window down, potentially due to a more refined temporal prediction and improved auditory-motor synchronisation (as suggested by the data of anisochrony and sensorimotor synchronisation tasks). The presently proposed paradigm provides a simple test to estimate the precision of this prediction. Musicians' temporal predictions were more precise but there were no reliable differences between pianists and brass players. The thresholds correlated with a purely auditory threshold measure requiring the detection of a temporal irregularity in an otherwise isochronous sound sequence. Furthermore, they correlated with sensorimotor synchronisation performance. This suggests that musical training improves a set of auditory and auditory-motor capacities. These capacities are used together to generate temporal predictions about the sensory consequences of our actions. The particular instrument as well as practice time has only a minor influence. This novel paradigm provides a simple test to estimate the strength of auditory-motor action-effect coupling that can readily be incorporated in a variety of studies investigating both healthy and patient populations.

It might be argued that the experimental setup of this study involves an inherent delay between the keystroke and the sound. Possibly, musicians who were exquisitely sensitive to delays considered even the shortest possible latency in our setup as asynchronous. However, if this were the case we would have expected participants to exhibit thresholds close to zero, which was not the case. Furthermore, as we have argued

above, the thresholds we found for musicians were comparable to those found in speech.

One limitation of our comparison between pianists and brass players is that the difference between those groups might have been reduced due to the fact that many brass players had some piano experience. This is not a bias in our sample, but reflects the reality of musical education in which musicians are encouraged to practice a secondary instrument, piano being a likely choice. Furthermore, we found no differences in a post-hoc comparison between those brass players with piano experience and those without.

Future studies could use the delay detection task to tap into temporal prediction capacities to investigate auditory-motor processing. The paradigm could also provide a precise quantification of temporal binding, which is the phenomenon that a person's self-generated sensory stimuli appear closer in time to the action that caused them than externally-generated sensory stimuli (Haggard, Clark, & Kalogeras, 2002). Furthermore, musicians with focal dystonia are shown to exhibit deficits in internal models (Lee, Furuya, Karst, & Altenmüller, 2013), which may be more precisely quantified with the present tools.

## References

- Alais, D., & Cass, J. (2010). Multisensory Perceptual Learning of Temporal Order: Audiovisual Learning Transfers to Vision but Not Audition. *PLoS ONE*, 5(6), e11283. doi:10.1371/journal.pone.0011283
- Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of cognitive neuroscience*, 21(4), 791–802. doi:10.1162/jocn.2009.21055
- Aschersleben, G. (2002). Temporal Control of Movements in Sensorimotor Synchronization. *Brain and Cognition*, 48(1), 66–79. doi:10.1006/brcg.2001.1304
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods*, 37(3), 379–384.
- Ben-Artzi, E., Fostick, L., & Babkoff, H. (2005). Deficits in temporal-order judgments in dyslexia: evidence from diotic stimuli differing spectrally and from dichotic stimuli differing only by perceived location. *Neuropsychologia*, 43(5), 714–723. doi:10.1016/j.neuropsychologia.2004.08.004
- Blakemore, S. J., Rees, G., & Frith, C. D. (1998). How do we predict the consequences of our actions? A functional imaging study. *Neuropsychologia*, 36(6), 521–529.
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7), 635–640. doi:10.1038/2870
- Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, perception & psychophysics*, 72(4), 1120–1129. doi:10.3758/APP.72.4.1120
- Drewing, K., Stenneken, P., Cole, J., Prinz, W., & Aschersleben, G. (2004). Timing of bimanual movements and deafferentation: implications for the role of sensory movement effects. *Experimental brain research*, 158(1), 50–57. doi:10.1007/s00221-004-1870-9
- Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: influence of the tempo and musical backgrounds of listeners. *Brain and Cognition*, 58(1), 133–147. doi:10.1016/j.bandc.2004.09.014
- Eliades, S. J., & Wang, X. (2003). Sensory-Motor Interaction in the Primate Auditory Cortex During Self-Initiated Vocalizations. *Journal of Neurophysiology*, 89(4), 2194–2207. doi:10.1152/jn.00627.2002

- Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Prozesse. III. *Pflugers Arkiv fur die gesammte Physiologie des Menschen und Thiere*, 11, 402–412.
- Fisher, N. I. (1995). *Statistical Analysis of Circular Data*. Cambridge University Press.
- Fostick, L., & Babkoff, H. (2013). Different Response Patterns Between Auditory Spectral and Spatial Temporal Order Judgment (TOJ). *Experimental Psychology*, 1(-1), 1–12. doi:10.1027/1618-3169/a000216
- Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., ... Driver, J. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex; a journal devoted to the study of the nervous system and behavior*. doi:10.1016/j.cortex.2013.03.006
- Frissen, I., Ziat, M., Champion, G., Hayward, V., & Guastavino, C. (2012). The effects of voluntary movements on auditory-haptic and haptic-haptic temporal order judgments. *Acta psychologica*, 141(2), 140–148. doi:10.1016/j.actpsy.2012.07.010
- Friston, K. (2012). Prediction, perception and agency. *International Journal of Psychophysiology*, 83(2), 248–252. doi:10.1016/j.ijpsycho.2011.11.014
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, 7(7), 773–778. doi:10.1038/nn1268
- García-Pérez, M. A., & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model. *Psychonomic bulletin & review*, 19(5), 820–846. doi:10.3758/s13423-012-0278-y
- Gaser, C., & Schlaug, G. (2003). Brain Structures Differ between Musicians and Non-Musicians. *The Journal of Neuroscience*, 23(27), 9240–9245.
- Gates, A., Bradshaw, J. L., & Nettleton, N. C. (1974). Effect of different delayed auditory feedback intervals on a music performance task. *Perception & Psychophysics*, 15(1), 21–25. doi:10.3758/BF03205822
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5(4), 382–385. doi:10.1038/nn827
- Helmuth, L. L., & Ivry, R. B. (1996). When two hands are better than one: reduced timing variability during bimanual movements. *Journal of Experimental Psychology. Human Perception and Performance*, 22(2), 278–293.
- Herholz, S. C., & Zatorre, R. J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, 76(3), 486–502. doi:10.1016/j.neuron.2012.10.011
- Hirsh, I. J., & Sherrick Jr, C. E. (1961). Perceived order in different sense modalities. *Journal of experimental psychology*, 62(5), 423.
- Hyde, K. L., & Peretz, I. (2004). Brains That Are Out of Tune but in Time. *Psychological Science*, 15(5), 356–360. doi:10.1111/j.0956-7976.2004.00683.x
- Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correlational analysis. *Acta Psychologica*, 60(2-3), 173–191.
- Keetels, M., & Vroomen, J. (2012). Exposure to delayed visual feedback of the hand changes motor-sensory synchrony perception. *Experimental brain research*, 219(4), 431–440. doi:10.1007/s00221-012-3081-0
- Kuling, I. A., van Eijk, R. L. J., Juola, J. F., & Kohlrausch, A. (2012). Effects of stimulus duration on audio-visual synchrony perception. *Experimental brain research*, 221(4), 403–412. doi:10.1007/s00221-012-3182-9
- Lee, A., Furuya, S., Karst, M., & Altenmüller, E. (2013). Alteration in forward model prediction of sensory outcome of motor action in focal hand dystonia. *Frontiers in Human Neuroscience*, 7, 172. doi:10.3389/fnhum.2013.00172
- Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral cortex*, 15(3), 299–302. doi:10.1093/cercor/bhh131
- Pfordresher, P. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology. Human Perception and Performance*, 29(5), 949–964. doi:10.1037/0096-1523.29.5.949
- Pfordresher, P., & Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research*, 66(1), 71–79.

- Powers, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *The Journal of neuroscience*, *29*(39), 12265–12274. doi:10.1523/JNEUROSCI.3501-09.2009
- Repp, B. H. (2004). On the nature of phase attraction in sensorimotor synchronization with interleaved auditory sequences. *Human movement science*, *23*(3-4), 389–413. doi:10.1016/j.humov.2004.08.014
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992.
- Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, 1–50. doi:10.3758/s13423-012-0371-2
- Rohde, M., & Ernst, M. O. (2012). To lead and to lag - forward and backward recalibration of perceived visuo-motor simultaneity. *Frontiers in psychology*, *3*, 599. doi:10.3389/fpsyg.2012.00599
- Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental brain research*, *227*(2), 249–261. doi:10.1007/s00221-013-3507-3
- Stuart, A., Kalinowski, J., Rastatter, M. P., & Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America*, *111*(5 Pt 1), 2237–2241.
- Sugano, Y., Keetels, M., & Vroomen, J. (2010). Adaptation to motor-visual and motor-auditory temporal lags transfer across modalities. *Experimental brain research*, *201*(3), 393–399. doi:10.1007/s00221-009-2047-3
- Szymaszek, A., Szélag, E., & Sliwowska, M. (2006). Auditory perception of temporal order in humans: The effect of age, gender, listener practice and stimulus presentation mode. *Neuroscience Letters*, *403*(1–2), 190–194. doi:10.1016/j.neulet.2006.04.062
- Tanaka, A., Asakawa, K., & Imai, H. (2011). The change in perceptual synchrony between auditory and visual speech after exposure to asynchronous speech. *Neuroreport*, *22*(14), 684–688. doi:10.1097/WNR.0b013e32834a2724
- Tillmann, B., Stevens, C., & Keller, P. E. (2011). Learning of timing patterns and the development of temporal expectations. *Psychological Research*, *75*(3), 243–258. doi:10.1007/s00426-010-0302-7
- Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2008). Audiovisual temporal adaptation of speech: temporal order versus simultaneity judgments. *Experimental brain research*, *185*(3), 521–529. doi:10.1007/s00221-007-1168-9
- Weiss, K., & Scharlau, I. (2011). Simultaneity and temporal order perception: Different sides of the same coin? Evidence from a visual prior-entry study. *Quarterly journal of experimental psychology (2006)*, *64*(2), 394–416. doi:10.1080/17470218.2010.495783
- Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, *14*(1), 5–12. doi:10.3758/BF03198607
- Yamamoto, K., & Kawabata, H. (2011). Temporal recalibration in vocalization induced by adaptation of delayed auditory feedback. *PloS one*, *6*(12), e29414. doi:10.1371/journal.pone.0029414
- Yamamoto, S., Miyazaki, M., Iwano, T., & Kitazawa, S. (2012). Bayesian calibration of simultaneity in audiovisual temporal order judgments. *PloS one*, *7*(7), e40379. doi:10.1371/journal.pone.0040379
- Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin*, *60*(3), 213–232. doi:10.1037/h0044155
- Yee, W., Holleran, S., & Jones, M. R. (1994). Sensitivity to event timing in regular and irregular sequences: influences of musical skill. *Perception & psychophysics*, *56*(4), 461–471.
- Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental brain research*, *152*(2), 198–210. doi:10.1007/s00221-003-1536-z

## Acknowledgements

This work was supported by the EBRAMUS, European Brain and Music Ph.D. Grant (ITN MC FP7, GA 238157). The team "Auditory cognition and psychoacoustics" is part of the LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-60). We are greatly indebted

to research assistant Phuong Mai Tran for implementing the experiment.

Sounds on time - Learning in musical novices

### **II.3. Auditory feedback benefits short term motor learning**

## Auditory feedback benefits short term motor learning

van Vugt FT<sup>1,2\*</sup>, Tillmann B<sup>2</sup>

<sup>1</sup>Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, CNRS-UMR 5292, INSERM U1028, University Lyon-1, Lyon, France

<sup>2</sup> Institute of Music Physiology and Musicians' Medicine, University of Music, Drama and Media, Hannover, Germany

### Abstract

*Introduction.* Humans learn skills by learning from their errors. The brain has been shown to use visual and proprioceptive feedback to update motor commands. However, it remains unclear whether the brain can also use auditory information such as, for example, when we learn to play music. Here, we ask whether participants (1) can benefit from auditory feedback, (2) whether this benefit depends on the time-coupling between the movement and the resulting sound, and (3) whether the learning in a particular feedback condition generalises to sequences other than those that were learned.

*Method.* Healthy non-musician participants learned to tap a sequence of seven numbers over the course of two days. Instructions were to tap as regularly as possible, keeping a constant speed. On the second day, they learned a second, novel sequence which they had not practised before. Participants were divided into three groups. In the sound group, participants heard a tone immediately every time they pressed a key. In the mute group, no keystroke-triggered sound was presented, but instead a continuous noise that masked their keystrokes. The jitter group received sounds after the keystroke with a random delay of 10-190 msec.

*Results.* Participants in the sound group were able to improve in tapping regularity whereas the mute and jitter group remained at the same level of irregularity. Furthermore, the sound group maintained their tapping regularity when they tapped the new sequence. After the learning, participants in the sound group tapped closer to the beat.

*Discussion.* Auditory feedback can be used by the brain in learning movement regularity. Our results indicate that the time-lock between movement and sound is crucial.

Furthermore, sound-supported learning is sequence-unspecific, providing a potential for rehabilitation applications.

### 3.1 Introduction

In a constantly changing environment, it is essential to be able to adapt by changing the way we behave. For example, we need to be able to adjust our grasping of a coffee cup to avoid obstacles that may or may not be there and we need to change our motor commands based on whether our muscles are fatigued or not. Similarly, a musician may be asked to play a melody with various rhythms, requiring control over the timing of individual movements. Correction of our actions happens both through on-line, fast correction of motor errors as well as through trial-to-trial updating of the motor program, enabling learning of complex skills such as mastering a musical instrument (Doyon, 2008; Wolpert, Diedrichsen, & Flanagan, 2011). One commonly held view is that motor learning occurs through a mechanism of error-correction (Shadmehr & Krakauer, 2008; Shadmehr, Smith, & Krakauer, 2010). Schematically, it can be thought of to proceed according to the following steps: (a) executing an action on the basis of our current motor program (Wolpert, Ghahramani, & Jordan, 1995; Wolpert, Miall, & Kawato, 1998), (b) perceiving the result of the action, (c) comparing this result to some desired template, (d) updating our internal representation on the basis of the perceived error. These steps are then repeated to form a chain of "cycles" or "iterations". In this way, our actions will usually converge to the template. The dynamics of such learning processes have been studied in the laboratory by making participants perform repetitive reaching movements in a force field. Over time, such reaching movements improve dramatically (e.g., Nezafat, Shadmehr, & Holcomb, 2001).

What kinds of feedback does the brain use for motor learning? For example, it is known that visual (Bernier, Chua, Bard, & Franks, 2006) and proprioceptive (Wong, Kistemaker, Chin, & Gribble, 2012) information play a large role in correcting ongoing movements, but also in planning fast, feed-forward controlled grasping movements (Ghez, Gordon, Ghilardi, & Sainburg, 1995). Also, tactile feedback seems crucial for online error correction in a tapping synchronisation paradigm (Aschersleben, Gehrke, & Prinz, 2001). Presumably, feedback from multiple modalities is used and integrated by weighing the

reliability of each modality (Kording & Wolpert, 2004).

Considerably less is known about the use of auditory feedback in motor learning. Can it be used to update motor programs? Investigations into the influence of auditory feedback have looked at sequence learning, timing coordination patterns, or music performance, which we will discuss in turn here below.

Firstly, motor sequence learning has been investigated with the serial-reaction-time paradigm (SRT) (Nissen & Bullemer, 1987). In this paradigm, participants are instructed to respond as quickly as possible to the appearance of one of various possible visuo-spatial cues by pressing the associated button. Unknown to participants, the sequence of appearance of the cues follows a fixed pattern. It is routinely found in SRT studies that participants' reaction time decrease over various blocks of training in this sequence and increase when the cues start following another sequence of appearance. Does sensory information influence learning? This question has been addressed in two ways: by adding sensory effects (tones) to participants' keystrokes and by investigating sensory learning of the cue sequence. Adding auditory effects to participants' keystrokes boosted learning. When participants received task-irrelevant tones at every keystroke, they exhibited greater learning (Hoffmann, Sebal, & Stöcker, 2001). This auditory benefit is only present if the pitch of the feedback tones is aligned with the left-to-right spatial alignment of the keys (Hoffmann et al., 2001; Stöcker, Sebal, & Hoffmann, 2003). The explanation for the boosted learning is that participants use the sensory effects of their actions as a control signal in the selection of their response. In particular, it has been suggested that sensory feedback influences the action initiation (response selection) stage but not the action execution stage (Kunde, Koch, & Hoffmann, 2004). Next, sensory learning has been proposed to form a part of the observed SRT learning. The fact that participants respond faster to cues in the course of the experiment may be due to the cue being predicted and therefore the response being primed. That is, the motor improvement may have an important sensory learning component (Remillard, 2003). By properly dissociating motor and sensory learning, it was shown that sensory learning happens but only for sequences of limited complexity (Deroost & Soetens, 2006). The addition of motor responses to participants' task enabled them to learn more complex sequences. However, a musician population was shown to be faster at the outset (when no learning had taken place) when they received auditory feedback tones

relative to when they received no sounds (Conde, Altenmüller, Villringer, & Ragert, 2012). In the course of this experiment, this difference gradually decreased, showing that auditory feedback had no influence on the *learning* process but only on the baseline performance.

Secondly, the influence of auditory feedback has been investigated in timing-coordination tasks. For instance, Ronsse et al., (2011) had participants learn to perform a 90deg out-of-phase bimanual cyclic wrist movement. One group received visual feedback in the form of a two-dimensional figure that showed their movement. A second group received auditory feedback in the form of sounds that marked the timing of particular points in their trajectories. The auditory group performed identically on a retention test without auditory feedback, but the visual group is severely impaired when visual feedback is removed. This suggested that the visual group had become dependent on the feedback whereas the auditory group had not. This is supported by the neuroimaging observation of remnants of visual activity even in the no-feedback retention test. The *guidance hypothesis* (Salmoni, Schmidt, & Walter, 1984) postulates that participants may become dependent on feedback and therefore performance would return to baseline once the feedback is removed.

It has also been shown that auditory feedback in the form of modulated pink noise can be used by participants in a motor performance task that required them to adapt on-line to disruptions in a force field (Oscari, Secoli, Avanzini, Rosati, & Reinkensmeyer, 2012). The response to auditory feedback was largely comparable to that for visual feedback. *Sonification* is a paradigm in sport sciences where movements are mapped to sounds in real-time. The resulting sounds has been employed as a method to improve performance in professional rowers (Schaffert, Mattes, & Effenberg, 2011).

Thirdly, the role of auditory feedback in motor performance has been investigated through expert performance. Various studies have investigated the effects of auditory feedback deprivation and delay on movement production in professional musicians or in speech production (as we are all experts in speech). A robust finding is that depriving pianists of auditory feedback influenced performance only marginally (Repp, 1999a) or not at all (Finney, 1997; Finney & Palmer, 2003; Gates, Bradshaw, & Nettleton, 1974). However, tones emitted with a constant delay after key press severely deteriorated performance (Gates & Bradshaw, 1974; Pfordresher, 2003; Pfordresher & Palmer, 2002). In speech

production, findings are similar. Deprivation of auditory feedback such as in post-lingually deaf individuals gradually deteriorates their speech production (Waldstein, 1990). Similarly, delayed auditory feedback severely disrupts speech production (Stuart, Kalinowski, Rastatter, & Lynch, 2002; Yates, 1963) as do shifts in the sound spectrum (Stuart, Frazier, Kalinowski, & Vos, 2008). These findings show that auditory feedback can affect motor performance. However, it might be argued that these results cannot be extrapolated to motor learning and to nonmusicians as they are based on a particular population of highly trained participants. In the case of professional musicians, the movements corresponding to their playing the instrument have been executed thousands or millions of times (Münste, Altenmüller, & Jäncke, 2002). Movements are therefore likely stored in motor programs that can be effected through feed-forward mechanisms, that is, without relying on auditory feedback anymore (Kawato, 1999; Lashley, 1951). Similarly, we are all experts in speaking. In other words, these studies investigated participants who already have established detailed internal representations, crystallised in stable motor programs, to find out whether these representations are disrupted by unexpected auditory feedback.

In sum, although several studies have investigated auditory feedback in motor processes, the surprisingly simple question whether auditory feedback improves motor performance relative to an absence of such feedback remains open. Here, we asked whether participants receiving auditory feedback are able to improve their movement more than those who do not. Furthermore, if so, can this feedback still be used when its association to movement is degraded, notably through introduction of jittered delay between the movement and the resulting sound? In order to quantify movement quality, we propose to investigate the smoothness of motor sequence tapping. The measure is used in previous studies on nonmusicians: Lappe, Herholz, Trainor, & Pantev (2008) observed that practice of an isochronous melody improves not only the speed of reproduction but also the evenness of playing, as measured by the standard deviations of the inter-note-intervals. Kuriyama and colleagues have participants learn to tap a simple sequence of numbers. They found that overnight the tapping intervals that were longest (the “problem points”) became shorter whereas the short intervals stay the same (Kuriyama et al., 2004). This suggests that overnight the sequence production will have become more smooth, although the authors

did not quantify the smoothness directly. Also, Keller & Koch (2006) have nonmusician participants respond in three isochronous taps continuing a visually established rhythm. They find that the produced intervals deviate less from the given tempo when the keys are mapped to task-irrelevant tones increasing pitch from left to right (congruent mapping), than when this mapping is reversed or scrambled (incongruent mapping). This finding reveals that if auditory feedback is to be beneficial, the key-to-tone mapping should be compatible with our expectations.

Here we use a simple motor sequence learning task in which participants repetitively tap a sequence of digits on a response button-box. The crucial manipulation is that some participants receive auditory feedback as a result of their keystrokes, either simultaneously or jittered in time, and others do not. Using this paradigm, we address the following questions. First, can auditory feedback be used to improve tapping evenness (smoothness) in a motor sequence production task? We hypothesised that participants could use the sound evoked by the keystrokes as a feedback signal to adjust their timing of the sequence. In this way, we expected that the group who received sounds would improve in tapping regularity more than a group who did not. Secondly, does the feedback sound need to be time-locked to the movement? If participants use sound as a feedback signal, their learning should be slower if the time-lock between the sound and movement was distorted. We therefore hypothesised that a group for whom the keystroke-evoked sounds were jittered in time would show less improvement in tapping regularity. Thirdly, how is performance affected when the auditory feedback is no longer present? Furthermore, does feedback influence whether the learned skills carry over through different sequences of movements? We addressed this latter question by enabling participants to learn to tap a particular sequence of keystrokes and then testing them on a second sequence under the same conditions.

We performed a motor learning experience in which participants learned to tap two sequences of seven digits over the course of two days using the index, middle, ring and little finger of their right hand. The participants' task was to tap the sequence correctly and with the speed and regularity indicated by a visual metronome in the form of a yellow bar indicating the items in the sequence prior to each trial. Practically, we asked participants to

tap a movement sequence at roughly the same speed, and to make their tapping as regularly as possible (instead of becoming faster). In order to enable participants to keep the same speed but avoid having to synchronise with an external stimulus, we introduce the (same) tempo prior each trial. In order to avoid the confounds arising from an auditory pacing signal, we used a visual “metronome” (Hove & Keller, 2010; Hove, Spivey, & Krumhansl, 2010) in the form of a bar that indicated the sequence items to be tapped on the screen prior to each trial.

The experiment spanned two days, during the first of which participants trained in one digit sequence that was assigned to them (primary sequence). The second day the participants continued with the first digit sequence and were then introduced to a novel sequence (secondary sequence). In addition to the sequence tapping training, participants completed basic auditory temporal change detection tasks (anisochny), keystroke-sound delay detection and synchronisation-continuation tapping (for details see below). The key experimental manipulation in this study related to the sound that was associated with each keystroke. Participants were divided into three groups. In the normal group, each keystroke triggered a (woodblock) sound through participants headphones, which occurred immediately upon the keystroke. In the mute group, participants' keystrokes did not trigger any sounds, and instead the headphones produced a continuous white noise (at a comfortable loudness level) to mask any residual mechanical keystroke sounds. In the jitter group, the keystrokes triggered a sound, but this sound was produced with a delay between 10 and 190 msec. The delay amount was sampled from a flat distribution (i.e., all values between 10 and 190 msec were equally likely) for each keystroke anew. Participants in all groups were instructed to ignore any sounds and care was taken to omit references to parallels with music in the instructions. For example, instead of instructing participants to copy the “metronome tempo” we instructed them to copy the “speed indicated by the bar”. Furthermore, we decided that all keys mapped to the same sound so that there were no different pitches for the different keys. In this way, we eliminated the possibility that participants used the sounds for response selection (such as in Hoffmann et al, 2001) instead of for response execution or timing.

In order to test participants' basic auditory temporal precision, participants performed the anisochny detection task (Ehrlé & Samson, 2005; Hyde & Peretz, 2004) that

measured their threshold for detecting a slightly late note in an otherwise isochronous five-tone sequence. Furthermore, in order to establish participants' temporal precision for the prediction of an auditory effect of their keystrokes, we used a delay detection task. In this task, we measured participants' sensitivity to delays between motor (keystroke) and auditory (tone) events. That is, we established from which delay onwards participants noticed that the tone came after the keystroke instead of immediately. Finally, in order to measure participants' auditory-motor synchronisation capabilities, they performed a synchronisation-continuation tapping paradigm (Tillmann, Stevens, & Keller, 2011; Wing & Kristofferson, 1973).

## 3.2 Methods

### 3.2.a Participants

Thirty-six right-handed non-musicians participated in the study. One participant was not able to press the individual keys independently and was therefore replaced by a different participant. Table 2 presents general characteristics of the groups. In that table, “marginal” refers to the summary for all groups combined; numbers are reported as mean (SD). We reported statistical comparisons of the three groups using Kruskal-Wallis tests or  $\chi^2$  for categorical distributions.

### 3.2.b Materials

#### i. General setup

The participants were seated and rested their arm comfortably on a table with the right hand index to little fingers placed on four keys of a custom button box. The participants' fingers rested on the keys with a little cushion of soft cloth in between. A removable shield was suspended above the button box so that the participants could not see their fingers or the keys. The participants were wearing headphones connected to the computer sound card which presented the auditory stimuli but blocked outside sounds.

The auditory white noise stimulus was generated using the program Audacity and saved as a wave file. The woodblock sound that sounded at each keystroke was 63 msec in duration and was chosen for its sharp onset (maximising temporal localisation) and nevertheless aesthetically pleasing sound. It was saved as a wave file and played using the experimental software.

The button box was custom made with four Cherry MX MX1A-11NN key switches and four key caps originating from the functional key row of the Diatec Filco set placed over them. In order to determine the optimal spatial arrangement for the buttons, we asked 6 scientific collaborators from our lab (3 male, 3 female) to put their fingers on a flat surface in the most comfortable and natural way. We marked the positions of the four fingers, averaged the distance matrices and transformed this average distance matrix back into 2 dimensional coordinates using MDS. We connected the MX keys through small electrical circuits to a commercially available button box (ioLab Systems, Inc.) that registered keystrokes and communicates these, timestamped, through serial USB interface with the Python script on Windows running pygame for visual presentation. The computer was a Dell Precision M6500 laptop running Windows 7 and interfacing with the box through HDI protocols in a python script. The sounds (woodblock or white noise) were presented through Sennheiser HD250 linear II headphones.

## ii. Sequences

Each participant was trained on two sequences of 7 digits (1-4), called *primary* and *secondary* sequence. Participants tapped the numbers 1 to 4 with the index (1) to little finger (4). The sequences had the same formal properties and were counterbalanced between the groups and between serving as primary and secondary sequence. We generated twelve different sequences according to the criteria below.

We designed our sequences to be of similar difficulty and give rise to major chunking, which has previously been shown to lead to timing deviations (Koch & Hoffmann, 2000). The participants' task, then, is to make their timing as regular as possible. Our pool of stimuli are 7-item sequences from the alphabet {1,2,3,4}. All numbers occur at least once, and no number more than twice. This gives  $4 \cdot 7! / 2 \cdot 2 = 2520$  possible basis sequences. We selected from those the sequences that have the following properties. Firstly,

sequences should contain exactly one immediate repetition (i.e.  $xx$  for  $x$  in  $\{1,2,3,4\}$ ) (e.g., Povel & Collard, 1982). Secondly, they should contain one trigram running upward or downward (e.g., 234 or 321) – because runs create very clear “chunks” in the sequence (Koch & Hoffmann, 2000). Thirdly, the sequence should contain at least two so-called ambiguous letters (Cohen, Ivry, & Keele, 1990). An ambiguous letter is a letter that occurs multiple times in a sequence and that on different occasions is followed by different letters (e.g., 3 is ambiguous in the sequence 3431). Technically, a letter  $x$  from the alphabet is ambiguous if the sequence contains  $xi$  and  $xj$  for some different  $i, j$  in the alphabet. Fourthly, the sequence should not contain repeating trigrams. Actually 276 sequences satisfy these constraints which amounted to 10.95% of all possible basis sequences. We randomly chose twelve sequences that satisfy the parameters mentioned above. Examples of such sequences are: 1234321, 1321442 (see supplementary materials for a complete listing). The set of sequences was then divided into six pairs. For each group (normal, jitter, mute) and each sequence pair  $(a, b)$  one participant had  $a$  as primary sequence and  $b$  as secondary sequence, and another participant  $b$  as primary and  $a$  as secondary sequence. In this way, the participants in each of the three groups, taken together, practiced the same sequences, and each sequence was equally often primary and secondary sequence.

### iii. Anisochrony task

Stimuli used in the anisochrony task consisted of a sequence of five isochronous sine wave tones of 100ms presented with an inter-onset-interval (IOI) of 350ms. In some trials, the fourth tone was delayed by a certain amount but the fifth tone was always on time (Ehrlé & Samson, 2005; Hyde & Peretz, 2004). That is, when the tone was delayed by an amount  $d$ , the third interval was longer by  $d$  msec and the fourth interval was shorter by  $d$  msec. Stimuli were presented through the same headphones used throughout the experiment, set to a comfortable sound level that was kept constant across all participants. Participants interacted with a python-pygame-based graphical user interface that showed the written instructions and collected the participant's button press responses from the laptop keyboard on each trial. In addition, instructions were given orally as well.

**iv. Delay detection task**

For the delay detection task, we used the same button box and headphones that were used during the sequence learning experiment. The woodblock sound that was presented on each trial was the same as that used during the sequence learning part.

**v. Synchronisation-continuation tapping task**

For the synchronisation-continuation task, we used the same button box that was used during the sequence learning experiment as well as the same headphones. The synchronisation stimulus was generated offline as follows and saved to a wave file. First, we created 4 finger snap sounds with an inter-onset-interval of 300 msec. Then followed 30 instances of the woodblock sound (the same as used during the learning part of the experiment) with an inter-onset-interval (IOI) of 600 msec. This was followed by a silence of 30\*600 msec (the equivalent of 30 more taps). Finally, a high-pitched gong sound was used to signal the end of the trial. The sounds were played using a custom developed python experimental script, which also communicated through a HID-USB interface with the button box to register the responses.

### **3.2.c Procedure**

Participants came into the lab on two adjacent days. In the beginning of the session on the first day, participants completed a brief questionnaire that established their handedness (according to the 10-item short form of the Edinburgh Handedness Inventory) and how much time they spend speed typing or phone messaging (as we suspected these would influence their sequence learning) and to verify that they had little or no musical experience. They then completed the sequence learning blocks for that day (more details below). The overall first day session lasted slightly less than an hour. On the second day, they first continued the sequence training blocks (details below). Then, participants completed the auditory and motor tests (anisochrony, delay detection and synchronisation-continuation tapping, see above). Finally, participants filled out a debriefing part of the questionnaire. This session on the second day lasted about 75 minutes. Participants received a nominal payment for their participation.

**i. Sequence tapping**

Participants were asked to tap a sequence of adjacent button positions (1234321), referred to hereafter as a scale sequence, prior to the sequence learning on the first day and after sequence learning on the second day. In one block of 25 passes, they were asked to tap the scale at a speed of 300 msec between keystrokes. In a second scale block, their task was to tap the scale as fast as they could (see Table 1). In all scale blocks, all three groups received white noise over their headphones and no keystroke-triggered sounds, so as to ensure that conditions were strictly comparable between the groups.

After the scale sequence on the first day, participants engaged in the sequence learning part of the experiment. The primary sequence is trained during the first day and part of the second day, and the secondary sequence is trained only on day two for several blocks. Our aim was to have participants tap the learning sequence at roughly the same speed throughout the experiment and monitor their improvement in regularity. At the same time, we designed our task to be demanding by requiring participants to produce the sequence at relatively high speed (300 msec between keystrokes). In order to make this feasible and not discourage our participants, we designed several training blocks of several trials (see Table 1 for full details) during which, initially, the participants could try tapping the sequence at their own speed. Then, pacing was introduced but at a slower speed (500 msec), gradually speeding up to 300 msec. Finally, the participants' tapping speed was verified to not be 10% slower or faster than indicated (300 msec) and they received a notice to speed up or slow down otherwise. During the final warm-up block, participants needed to tap the sequence 5 times correctly and within the aforementioned speed window before the experiment continued. Importantly, in the warm-up blocks all three groups received white noise over their headphones and no keystroke-triggered sounds. This was done so as to ensure that no auditory-motor learning could happen during the warm-up. Only in the last warm-up block, participants in the normal and jitter group received keystroke-triggered sounds (either immediately or randomly delayed, respectively).

On the second day, after having completed the main training blocks for the primary sequence (after block 10) and for the secondary sequence (after block 3), participants were asked to complete one more block, this time without receiving any keystroke-triggered sound. That is, for the mute group the situation did not change, but the normal and jitter

groups no longer received keystroke-triggered sounds but instead heard white noise at a comfortable loudness level. For a complete overview of the order of the training blocks, as well as their details, please refer to Table 1.

## **ii. Procedure for one trial**

The seven digits of the sequence (primary, secondary or scale) were presented at once on the computer screen coded as numbers 1 (index) to 4 (little finger) (Povel & Collard, 1982). On each trial, the participant first passively watched a yellow bar indicate the items in the sequence one by one at the desired speed (300 msec IOI except during warmup blocks) (Figure 1). Once the yellow bar had indicated the last item of the sequence, the participants commenced tapping the sequence from the beginning. At the first keystroke, the yellow bar disappeared and the first item in the sequence was slightly greyed out to indicate that it had been tapped. Subsequent keystrokes (whether correct or not) caused the subsequent items in the sequence to be greyed out. The greying out was chosen to be very subtle so that it made clear the key was struck, but provide as little timing cues as possible. In this way, participants could not use the timing of the greying out as a timing cue in their motor learning. Note that the greying out was constant across the three conditions.

Furthermore, we decided to use only a single feedback tone (a woodblock sound) that did not differ between the different keys. In this way, we avoid that participants would use the pitch contour to guide their movements. Also, we filter out pitch-to-space mapping compatibility effects reported previously (Keller & Koch, 2006). Using a same sound or using different sounds has to our knowledge not been contrasted in a single study previously. Ullen & Bengtsson (2003) used a single (drum) sound in one experiment and various tones in a second, but the study allows no direct comparison because of the lack of strictly overlapping conditions between the two experiments.

## **iii. Anisochrony task**

On each trial, an isochronous sequence of five tones in was presented through their headphones. In some cases, the fourth tone was delayed (see materials). Participants' task was to respond whether they heard a change in the intervals or not. The participant threshold for detecting such temporal deviation was established adaptively using the

on-line maximum-likelihood procedure (MLP) (Green, 1993; Gu & Green, 1994). We used the Maximum Likelihood Procedure (MLP) algorithm to detect the threshold for the detection of the asynchrony between movement (keypress) and the tone. The algorithm is designed to adaptively select the stimulus level (tone delay) on each trial so as to converge to the participants' threshold. For each block, the algorithm outputs an estimate for the participant's threshold. The MLP algorithm briefly works as follows. A set of candidate psychometric curves are maintained in parallel and for each, the likelihood of the set of the participants' responses is calculated. The psychometric curve that makes the participant's responses maximally likely is used to determine the stimulus level (the delay between the keystroke and the sound) on the next trial. We defined 200 logistic psychophysical curves whose midpoints were linearly spread over the range from 0 to 200 msec delay (0% to 57% of the tone IOI) and these were crossed with the five false alarm rates (0, 10, 20, 30, 40%). The maximum likelihood procedure was implemented in python. We made our source code freely available online on <https://github.com/florisvanvugt/PythonMLP>.

Participants first performed 4 trials (2 with no delay and 2 with a delay of 200 msec) to make clear the difference between when the sound came immediately and when it was delayed. The participant received accuracy feedback about her answers during these practice trials. Next, they performed a block of 10 trials, starting at 200 msec delay but then using MLP to determine the stimulus levels of the following trials. If the procedure was clear, we continued with 3 experimental blocks of 36 trials. Each experimental block consisted of 36 trials containing 6 catch trials. Catch trials are trials on which the delay was always 0 msec (regardless of the delay that was suggested by the MLP algorithm). The function of catch trials is to prevent participants from always responding “delayed” (which would cause the MLP algorithm converge to a zero threshold). Catch trials were inserted randomly as follows: the first 12 trials contained 2 catch trials and the next 24 trials contained 4 catch trials.

#### **iv. Delay detection task**

In each trial of the delay detection task, participants pressed the key of the button box that corresponded to their index finger. After a specified delay, a sound was presented through their headphone. The participants' task was to respond verbally whether the sound

came immediately after the keystroke or whether there was a delay. This verbal response was entered into the experimental software by the experimenter. Crucially, participants were instructed to leave their finger on the key (instead of lifting it prior to the keystroke) so as to reduce the tactile timing information. Furthermore, they were required to keep their eyes closed during the keystroke.

We used the Maximum-Likelihood Procedure (MLP) algorithm to detect participants' threshold for detecting a delay between their keystroke and the sound. The basic procedure of the algorithm was the same as for the anisochrony task, but here the set of candidate psychometric curves was as follows. We used 600 candidate psychometric curves with midpoints linearly spread between 0 and 600ms delay and crossed by the five false alarm rates (0%, 10%, 20%, 30%, 40%). Hence, a total of 3000 candidate psychometric curves were used.

Participants first performed 4 trials (2 with no delay and 2 with a delay of 600ms) to make clear the difference between when the sound came immediately and when it was delayed. The participant received accuracy feedback about her answers during these practice trials. Next, they performed a block of 10 trials, starting at 600ms delay but then using MLP to determine the stimulus levels of the following trials. If the procedure was clear, we continued with 3 experimental blocks of 36 trials.

#### **v. Sensorimotor synchronisation task**

In each trial, participants tapped with their index finger on a flat surface along with the synchronisation stimulus (see materials). When the woodblock sounds stopped, participants were instructed to continue tapping at the same speed and regularity until the high-pitched sound signaled the end of the trial.

Day	Block	Sequence	Trials	Pacing	Auditory feedback
1	Scale	Scale	5	Unpaced free	All groups mute
1	Scale-300	Scale	25	300 msec	All groups mute
1	Scale-max	Scale	25	Unpaced maximal	All groups mute
1	Warmup-free	Primary	5	Unpaced free	All groups mute
1	Warmup-500/400/300	Primary	5/5/5	500/400/300 msec	All groups mute
1	Warmup-300-fixed	Primary	5 correct and in time	300 msec+	By group
1	Main block 1-7	Primary	25 each	300 msec	By group
2	Warmup-300	Primary	5	300 msec	By group
2	Warmup-300-fixed	Primary	5 correct and in time	300 msec+	By group
2	Main block 8-10	Primary	25 each	300 msec	By group
2	Mute block	Primary	25	300 msec	All groups mute
2	Warmup-free	Secondary	5	Unpaced free	All groups mute
2	Warmup-500/400/300	Secondary	5/5/5	500/400/300 msec	All groups mute
2	Warmup-300-fixed	Secondary	5 correct and in time	300 msec+	By group
2	Main block 1-3	Secondary	25 each	300 msec	By group
2	Mute block	Secondary	25	300 msec	All groups mute
2	Scale-300	Scale	25	300 msec	All groups mute
2	Scale-max	Scale	25	Unpaced maximal	All groups mute

*Table 1:* Listing of the blocks in the motor learning experiment. The sequence was either a scale (1234321), the primary assigned sequence or the secondary assigned sequence. *Pacing* indicates whether the trials were paced (by a visual metronome indicating the numbers one by one prior to the trial) ("+" indicates that the this pacing was then enforced by repeating trials that were more than 10% slower or faster than the target tempo). When *auditory feedback* is listed as “mute”, this indicates that in those blocks all participants heard only continuous white noise (instead of keystroke-triggered sounds), whereas “by group” indicated that the feedback condition relative to the assigned group was used, that is ‘normal’, ‘jittered’ or ‘mute’.

### 3.2.d Data Analysis

#### i. General analysis methods

The sequence tapping data were analysed as follows. We discarded sequences that contained errors (additional keystrokes) or omissions and then performed the following two-step outlier rejection procedure. First, we pooled the data from all participants and discarded inter-keystroke-intervals that exceeded 5 SD of the overall mean. This figure was deliberately chosen conservatively to discard intervals that corresponded to sporadic keystrokes that failed to register and were attempted again by the participant after a long pause; this was confirmed by visual inspection and discarded 0.5% of all keystrokes. In a second step, we separated the intervals by participant and block and then eliminated all keystrokes that were further than 3 SD from the mean for that participant's block. This resulted in an additional 1.7% of all intervals being discarded. The reason for using this two-step discarding procedure is that discarding outliers based on parametric statistics (mean and SD) is known to be highly susceptible to the outliers in the first place, because the outliers strongly bias the initial estimates of the mean and SD themselves. Using the procedure above, we were able to first eliminate far outliers due to unregistered keystrokes, and later trim the distributions depending on the learning.

The main dependent variable of interest was the tapping regularity. For each participant and each block, we calculated the standard deviation of the inter-keystroke-interval durations and expressed it as a percentage of the mean interval duration so as to compensate for possible speed differences. The larger this quantity, the more irregular the tapping is, and the smaller, the closer to isochrony. In addition to the main dependent variable of tapping regularity, we also report accuracy (number of correct trials) and speed (the mean inter-keystroke-interval).

We performed mixed design ANOVAs unless otherwise indicated. Mauchly's test for sphericity was always performed and whenever it was significant, we applied the Greenhauser-Geisser correction and, for brevity, immediately report the corrected p-value marked with  $p_{GG}$ . We report generalised effect sizes ( $\eta_G^2$ ) (Bakeman, 2005). Group comparisons were calculated using Tukey's HSD procedure.

**ii. Anisochrony and Delay Detection tasks**

The delay detection and anisochrony threshold tasks were analysed as follows. First, we discarded blocks that contained more than 30% incorrect catch trial responses (in which the delay or deviation was 0 msec). Secondly, we discarded blocks in which the threshold estimate had not properly converged towards the end of the block. This was tested by fitting a regression line to the last 10 trials in the block and discarding those blocks in which the slope of this line exceeded 2 msec/trial (for the delay detection task) or 1.18 %IOI/trial (for the anisochrony task). These slope cut-off points were chosen to confirm visual inspection of blocks that had not properly converged. Thirdly, we computed the average threshold estimate for the remaining blocks for each participant.

**iii. Synchronisation-continuation tapping**

Synchronisation-continuation tapping performance was analysed using linear and circular statistics. In the linear analysis, we calculated the time between each tap and its corresponding metronome click (in msec). For each block, we averaged these to yield the mean relative asynchrony (in msec) and calculated the standard deviation to yield the SD relative asynchrony (in msec). The mean relative asynchrony is a measure for how close participants tapped to the beat and the SD relative asynchrony is a measure of tapping precision (time-lock). In the circular analysis (Fisher, 1995), the timing of each tap was converted into a phase (between 0 and  $2\pi$ ) relative to the metronome onset. Based on these, we calculated the synchronisation vector, which is the average of all vectors with length 1 and the phase angle for that tap. The length of this vector (between 0 and 1) is a measure for the time-lock between the tap and the sound. In order to average these vectors within groups, we used Fisher's r-to-z transformation and obtained the corresponding z-scores.

For the continuation phase (when the metronome had stopped), we calculated the interval between taps (inter-tap-interval, ITI, in msec) and its standard deviation (SD ITI in msec). We then de-trended the continuation taps by fitting a regression line to the ITIs over time, reporting the slope of this line and taking the residual variability from this line. In this way, we compensated for the fact that participants tend to speed up or slow down (Drewing, Stenneken, Cole, Prinz, & Aschersleben, 2004; Helmuth & Ivry, 1996; Keele,

Pokorny, Corcos, & Ivry, 1985). The slope of this line fit indicated the tempo drift.

	Jitter	Mute	Normal	Marginal	Statistical Test
N	12	12	12	36	N/A
Age (years)	25.1 (4.54)	23.6 (1.78)	25.7 (5.61)	24.8 (4.26)	Kruskal-Wallis $\chi^2(2)=.37$ p=.83
Gender (female/male)	5/7	9/3	6/6	20/16	$\chi^2(2)=2.9$ p=.23
Handedness (Edinburgh Laterality Quotient%)	78.9 (22.91)	82.3 (25.30)	84.3 (15.02)	81.8 (21.03)	Kruskal-Wallis $\chi^2(2)=.47$ p=.79
Text messaging capacity (self-rating)	6.6 (2.07)	7.5 (1.38)	7.3 (1.78)	7.1 (1.76)	Kruskal-Wallis $\chi^2(2)=2.08$ p=.35
Keyboard typing capacity (self-rating)	6.7 (1.30)	7.5 (0.67)	7.1 (1.56)	7.1 (1.25)	Kruskal-Wallis $\chi^2(2)=2.52$ p=.28
Blind typing capacity (number of participants in the following categories: 10 fingers/ <10 fingers / none)	4/1/7	3/4/5	5/3/4	12/8/16	$\chi^2(4)=3.13$ p=.54
Use of video games per week (number of participants in the following categories: 0h/<1h/1-7h/>7h)	3/4/2/3	6/3/2/1	7/1/4/0	16/8/8/4	Kruskal-Wallis $\chi^2(2)=2.72$ p=.26
Use of computer keyboard per week (number of participants in the following categories: <1h/1-14h/>14h)	0/5/7	0/6/6	0/7/5	0/18/18	Kruskal-Wallis $\chi^2(2)=.65$ p=.72
Use of cell phone text messages per week (number of participants in the following categories: <1h/1-7h/>7h)	5/3/4	3/8/1	3/7/2	11/18/7	Kruskal-Wallis $\chi^2(2)=.07$ p=.96
Years of musical instruction (other than obligatory courses at school).	0.7 (1.23)	0.1 (0.29)	0.9 (1.28)	0.5 (1.06)	Kruskal-Wallis $\chi^2(2)=3.96$ , p= 0.14

Table 2: Participant characteristics. Values are reported as mean (SD) unless otherwise specified. “Marginal” indicates the totals for all participants combined. We report statistical comparison between the three groups using Kruskal-Wallis tests (when the values are ordered) or  $\chi^2$  testing (when they are not ordered).

### 3.3 Results

#### 3.3.a Sequence Learning

We investigated how the learning of the participants proceeded and how this differed between the groups.

In order to establish whether participants produced more correct sequences in the course of the first day, we performed a 6x3 ANOVA with block (six levels corresponding to the six blocks) as within-participant factor, group (normal, mute and jitter) as between-participant factor and number of produced sequence as dependent variable. Participants gradually tapped more correct sequences in each block, from 20.5 (SD 3.7) correct sequences in the first block (out of 25) to 22.7 (SD 2.9) on the last block of the first day [ $F(6,198)=5.15$ ,  $p_{GG}=.0002$ ,  $\eta_G^2=.03$ ]. There were no differences between the groups [ $F(2,33)=.28$ ,  $p=.76$ ] and no interaction [ $F(12,198)=1.25$ ,  $p=.25$ ]. Similarly, participants were initially slightly slower than the tempo indicated by the visual metronome but then speeded up over the course of the first day of training [ $F(7,231)=5.12$ ,  $p_{GG}=.002$ ,  $\eta_G^2=.04$ ]. They tapped at 350.1 (SD 48.6) msec average interval duration on the first block. Gradually, they speeded up to 326.5 (SD 54.1) msec average interval duration at the last block of day 1. There was no main effect of group [ $F(2,33)=.65$ ,  $p=.52$ ] nor an interaction with group [ $F(14,231)=1.05$ ,  $p_{GG}=.40$ ] (Figure 2 top).

We found that the two groups were not different in regularity during the last warmup block [ $F(2,33)=0.43$ ,  $p=.66$ ] nor during the first main block [ $F(2,33)=2.01$ ,  $p=.15$ ]. However, taking all the blocks of the first day, there was a trend for a main effect of group [ $F(2,33)=2.86$ ,  $p=.07$ ] as well as a trend for an effect of block [ $F(7,231)=2.28$ ,  $p=.03$ ,  $p_{GG}=.076$ ,  $\eta_G^2=.01$ ] and a significant interaction between group and block [ $F(14,231)=2.00$ ,  $p=.02$ ,  $p_{GG}=.06$ ,  $\eta_G^2=.02$ ]. Analysing the groups separately, we found that the mute [ $F(7,77)=.94$ ,  $p_{GG}=.42$ ] and jitter [ $F(7,77)=1.42$ ,  $p_{GG}=.25$ ] groups showed no changes in tapping irregularity between the blocks, but the normal group showed a main effect of block [ $F(7,77)=3.05$ ,  $p_{GG}=.047$ ] indicating that the tapping of the sequence became more regular for this group over the course of the first day.

On the second day, participants produced 22.8 (SD 2.5) correct sequences on each

block of the primary sequence. This did not differ between blocks [ $F(2,66)=2.17$ ,  $p_{GG}=.13$ ], between the groups [ $F(2,33)=1.39$ ,  $p=.26$ ] and there was no interaction [ $F(4,66)=1.69$ ,  $p_{GG}=.17$ ]. Furthermore, the tapping speed was now at 324.1 (SD 45.6) msec (still above the strictly indicated 300 msec) and did not differ between the groups [ $F(2,33)=1.30$ ,  $p=.28$ ], remained constant between the blocks [ $F(2,66)=1.27$ ,  $p=.29$ ] and there was no interaction [ $F(4,66)=0.49$ ,  $p=.74$ ] (Figure 2 top). In terms of tapping regularity, we found no main effect of block [ $F(2,66)=2.22$ ,  $p_{GG}=.12$ ] and no interaction between group and block [ $F(4,66)=.72$ ,  $p_{GG}=.56$ ]. However, there was a main effect of group [ $F(2,33)=3.19$ ,  $p=.05$ ,  $\eta_G^2=.15$ ]. Post-hoc Tukey HSD testing revealed that this effect was due to the normal group tapping more regular than the jitter group [ $p=.065$ ] but not the mute group [ $p=.13$ ] and the jitter and mute group were not different [ $p=.94$ ] (Figure 2 bottom).

For the secondary sequence, participants again tapped more correctly during the third block ( $M=22.4$ ,  $SD=2.2$  correct sequences) than the first block ( $M=20.6$ ,  $SD=3.8$  correct sequences) [ $F(2,66)=8.84$ ,  $p_{GG}=.0007$ ,  $\eta_G^2=.05$ ]. There was a trend for a main effect of group [ $F(2,33)=2.73$ ,  $p=.08$ ] (indicating that the normal group produced slightly more correct sequences than the other two groups) but no interaction between group and block [ $F(4,66)=1.58$ ,  $p=.19$ ]. In terms of speed, participants again achieved greater speed in the later blocks [ $F(2,66)=9.63$ ,  $p_{GG}=.0004$ ,  $\eta_G^2=.04$ ] but there was no main effect of group on speed [ $F(2,33)=.75$ ,  $p=.48$ ] and no interaction [ $F(4,66)=.85$ ,  $p=.50$ ]. In terms of irregularity, we found no effect of block [ $F(2,66)=.02$ ,  $p=.98$ ] and no interaction between group and block [ $F(4,66)=.68$ ,  $p=.60$ ]. However, there was a main effect of group [ $F(2,33)=7.10$ ,  $p=.003$ ,  $\eta_G^2=.27$ ]. Tukey HSD tests indicated that the normal group was more regular than the jitter [ $p=.008$ ] and the mute [ $p=.003$ ] groups, but the mute and jitter groups were equally (ir)regular [ $p=.92$ ] (Figure 2 bottom).

### 3.3.b Sequence switching

In order to investigate whether participants' sequence learning for the primary sequence transferred to the secondary, novel sequence, we compared the last block of the primary sequence and the first block of the secondary sequence. We furthermore included the first block of the primary sequence in the ANOVA, so as to compare the starting blocks of the two sequences. That is, we investigate a three-level factor block (block 1 and 10 of the

primary sequence and block 1 of the secondary sequence) as within-participants factor and group as between-participants factor.

In number of correct sequences, we found no main effect of group [ $F(2,33)=.87$ ,  $p=.42$ ] but a main effect of block [ $F(2,66)=12.36$ ,  $p<.001$ ,  $\eta_G^2=.10$ ] and a tendency for an interaction between group and block [ $F(4,66)=2.45$ ,  $p=.053$ ,  $\eta_G^2=.04$ ]. We proceeded to analyse the three groups separately using post-hoc Tukey contrasts. The normal group improved in accuracy between the first and last block of the primary sequence [ $p=.001$ ]. Further, the first block of the secondary sequence [ $p=.04$ ] had more correct sequences than the first block of the primary sequence. But the last block of the primary and first block of the secondary sequence were not different [ $p=.87$ ]. The mute group, however, showed only a trend for an improvement between the first and last block of the primary sequence [ $p=.06$ ] and then a sharp decline in accuracy with the secondary sequence [ $p=.0006$ ]. There was no difference between the first blocks of the two sequences [ $p=.49$ ]. The jitter group showed a decline in accuracy between the last block of the primary sequence and the secondary sequence [ $p=.045$ ] but no differences between the other pairs of blocks [ $p=1.00$  and  $p=.24$ ] except (Figure 3).

As for tapping speed, there was no main effect of group [ $F(2,33)=1.60$ ,  $p=.22$ ] and no interaction between group and block [ $F(4,66)=.60$ ,  $p=.73$ ]. However, there was a main effect of block [ $F(2,66)=11.87$ ,  $p<.001$ ,  $\eta_G^2=.11$ ]. Post-hoc Tukey contrasts revealed that this signaled a strong difference between the first and last primary sequence block [ $p<.0001$ ] and differences between the first block of the secondary sequence and the first [ $p=.04$ ] and last [ $p=.04$ ] blocks of the primary sequence. Finally, we turn to the tapping irregularity, which is the main dependent variable under investigation. We find a main effect of group [ $F(2,33)=4.80$ ,  $p=.01$ ,  $\eta_G^2=.19$ ], a main effect of block [ $F(2,66)=.605$ ,  $p=.004$ ,  $\eta_G^2=.03$ ] and an interaction between the two [ $F(4,66)=3.33$ ,  $p=.02$ ,  $\eta_G^2=.04$ ]. We therefore proceeded to compare the three blocks for each group individually. The normal group became more regular between the first and last block of the primary sequence [ $p=.01$ ] and between the first blocks of the two sequences [ $p=.0002$ ] but the last block of primary sequence was not different from the first block of the secondary [ $p=.77$ ]. The mute group only showed an improvement from the first and last primary block [ $p=.02$ ] but not between the other block pairs [ $p=.98$  and  $p=.26$ ]. Lastly, the jitter group showed no significant differences between

any block pairs [all  $p=1.0$ ] (Figure 3).

### 3.3.c Muting effects

What was the effect of eliminating the sound? We exposed both the normal and jitter groups to a “mute” block, in which they heard only white noise and no longer received the woodblock sound at each tap. This means they were in the same physical condition as the mute group, who also performed another block in those same conditions. Note that this manipulation happened twice during the experiment (both times on the second day): once after the last (tenth) block of the primary sequence, and once after the last (third) block of the secondary sequence. We performed a repeated-measures ANOVA with factors sequence (primary, secondary), condition (mute vs. sound) and with the between-subjects factor group (normal or jitter). We did not include the mute group in this analysis because they were in a mute condition throughout the experiment.

For the number of correct sequences, we found a main effect of condition (mute vs. sound) indicating that both groups tended to produce *less* correct sequences in the mute condition [ $F(1,22)=9.18$ ,  $p=.006$ ,  $\eta_G^2=.05$ ]. There were no significant main effects of group [ $F(1,22)=2.88$ ,  $p=.10$ ] nor of sequence [ $F(1,22)=.35$ ,  $p=.56$ ]. There were no two-way interactions [all  $F(1,22)<1.03$ ,  $p>.32$ ] and the three-way interaction between group, sequence and condition was not significant [ $F(1,22)=1.40$ ,  $p=.45$ ] (Figure 4 top).

For sequence production speed, we found an interaction effect of group (normal, jitter) and condition (sound vs. mute) [ $F(1,22)=4.91$ ,  $p=.037$ ,  $\eta_G^2=.02$ ]. Therefore, we continued to analyse the two groups separately. For the normal group, there were no effects of sequence [ $F(1,11)=1.20$ ,  $p=.30$ ] or condition [ $F(1,11)=2.40$ ,  $p=.15$ ] and no interaction between sequence and condition [ $F(1,11)=.24$ ,  $p=.64$ ]. For the jitter group there was no effect of sequence [ $F(1,11)=.99$ ,  $p=.34$ ]. However, there was a main effect of condition [ $F(1,11)=20.85$ ,  $p=.0008$ ,  $\eta_G^2=.12$ ]. This effect of condition indicated that the jitter group tapped faster during the mute block than during the sound block. The interaction between condition and sequence was not significant [ $F(1,11)=.02$ ,  $p=.88$ ] (Figure 4 middle).

For tapping regularity, we found a main effect of group [ $F(1,22)=5.95$ ,  $p=.02$ ,  $\eta_G^2=.18$ ] which indicated that the normal group tapped more regularly than the jitter group in all

four blocks under consideration here. There was no effect of sequence (primary vs. secondary) [ $F(1,22)=1.42, p=.25$ ] or condition [ $F(1,22)=.16, p=.69$ ]. There were no interaction effects [all  $F(1,22)<1.75, p>.20$ ] (Figure 4 bottom).

### 3.3.d Scale tapping

First, we will discuss the blocks in which participants were instructed to tap the sequence at the speed indicated by the metronome (“paced” prior to each trial), and secondly, we will discuss the subsequent maximal speed block.

For the paced scale blocks and the speeded scale blocks, we performed a 2x2 ANOVA with day (day one before the learning part; day two after the learning part) as within-participant factor and group as between-participant factors, and with number of correct scales (out of 25), speed or variability as dependent variable, respectively. As for accuracy of the paced scale blocks, we found no effect of group [ $F(2,33)=1.86, p=.17$ ] nor of day [ $F(1,33)=1.97, p=.17$ ]. There was a tendency for an interaction between group and block [ $F(2,33)=3.19, p=.054, \eta^2=.06$ ], revealing that the jitter group produced less correct scales on the second day than on the first day [ $F(1,11)=8.56, p=.01, \eta^2=.16$ ] whereas there was no difference for the normal [ $F(1,11)=.20, p=.67$ ] and mute [ $F(1,11)=.03, p=.86$ ] groups (Figure 5). As for speed, there were no speed differences between the groups [ $F(2,33)=.97, p=.39$ ] and no interaction between day and group [ $F(2,33)=1.66, p=.20$ ]. However, there was a main effect of day [ $F(1,33)=43.62, p<.0001, \eta^2=.29$ ] indicating that participants approached the target IOI (300 msec) more closely on the second day than on the first day (Figure 5). In terms of tapping irregularity, there was a main effect of group [ $F(2,33)=4.20, p=.02, \eta^2=.14$ ] which revealed that the normal group was more regular in tapping than the mute and jitter groups on the second day [ $F(2,33)=4.08, p=.03, \eta^2=.20$ ] but not on the first day [ $F(2,33)=1.77, p=.19$ ] (Figure 5).

In the speeded scale blocks, participants produced more correct sequences on the second day [ $F(1,33)=9.44, p=.004, \eta^2=.05$ ], but there was no effect of group [ $F(2,33)=.39, p=.68$ ] and no interaction between group and block [ $F(2,33)=2.02, p=.15$ ]. On the second day participants were faster [ $F(1,33)=76.87, p<<.0001, \eta^2=.14$ ], but again no differences between the groups [ $F(2,33)=.26, p=.78$ ] or interaction [ $F(2,33)=1.97, p=.15$ ] appeared. In

parallel, participants showed a marked decrease in regularity [ $F(1,33)=23.68$ ,  $p<.0001$ ,  $\eta_G^2=.14$ ] on the second day relative to the first day. There were no differences between the groups [ $F(2,33)=.28$ ,  $p=.76$ ] and no interaction [ $F(2,33)=.81$ ,  $p=.45$ ]. This suggests that participants favoured speed over accuracy on the second day (Figure 5).

### **3.3.e Anisochrony detection task**

First, we discarded blocks with more than 30% errors in catch trials (8.3% of all blocks). Furthermore, an additional 2.7% of all blocks was discarded because of non-convergence. At this point, for one participant all three blocks had been discarded, and therefore this participant was eliminated from further analysis. For the other participants, we took the mean of the remaining 2.7 (0.6) blocks as estimate of their psychophysical threshold. We eliminated one other participant (from the jitter group) whose threshold was more than 3 SD away from the mean of all participants taken together. We found no difference of anisochrony threshold between the groups [ $F(2,31)=.27$ ,  $p=.76$ ], even when we included the outlier participant from the jitter group [ $F(2,32)=1.10$ ,  $p=.34$ ] (Table 3).

### **3.3.f Delay detection task**

We discarded blocks with more than 30% error catch trials (5.6% of all blocks). Furthermore, 18.5% of all blocks were discarded due to non-convergence. All participants had at least one block remaining, and we took the average of the thresholds of the 2.3 (SD 0.7) remaining blocks per participant. One participant (again in the jitter group, albeit a different participant from the one discarded in the anisochrony case) had a threshold more than 3 SD away from the overall mean, and was therefore discarded. There was no difference in thresholds between the groups [ $F(2,32)=1.17$ ,  $p=.32$ ], even when we reinserted the discarded participant [ $F(2,33)=2.09$ ,  $p=.14$ ] (Table 3).

### **3.3.g Synchronisation-Continuation Tapping**

One participant (from the jitter group) had to be discarded because he/she was not able to synchronise with the metronome in any of the trials. For one participant two out of three blocks exhibited many missed taps (due to a technical problem). These blocks were

discarded but the remaining block was included in the analysis. For synchronisation performance, circular statistics indicated that there were no differences between the groups in number of taps registered during this phase nor in synchronisation accuracy (Table 3). In order to investigate whether participants tapped closer to the beat, we verified that the synchronisation angles for the participants in each group could be approximated by a von Mises distribution [Watson's test statistic .05-.09, all  $p > .1$ ]. We then performed a circular ANOVA with between-participants factor group (normal, mute or jitter) and the synchronisation angle as a dependent variable. We found that participants the normal group tapped closer to the beat than the other two groups [circular ANOVA  $F(2,32)=4.08$ ,  $p=.026$ ,  $\eta^2=.18$ ]. Paired comparisons revealed that the normal group was closer to the beat than the mute group [ $F(1,22)=7.85$ ,  $p=.01$ ] and the jitter group [ $F(1,21)=4.13$ ,  $p=.055$ ] but the jitter and mute group were not different [ $F(1,21)=.51$ ,  $p=.48$ ] (Figure 6). In order to investigate whether sequence learning performance could be linked to this synchronisation difference, we computed the average sequence tapping regularity for the last three (non-mute) sequence tapping blocks. In these blocks, participants tapped the secondary sequence. The tapping regularity correlated negatively with the tap-click distance during synchronisation-continuation tapping [Spearman  $\rho(34)=-.42$ ,  $S=10156$ ,  $p=.01$ ]. This indicated that those who tapped the sequence more regularly also tapped closer to the beat. In order to analyse the continuation phase, we corrected for tempo drifts by linear de-trending: we fitted a line and analysing the residuals from this line (Drewing et al., 2004; Helmuth & Ivry, 1996; Keele et al., 1985). The slope of the line indicated the tempo drift and was not different between groups nor was the residual variability (Table 3).

	<b>Jitter</b>	<b>Mute</b>	<b>Normal</b>
Anisochrony Threshold (% of IOI)	12.3 (4.8)	11.3 (5.9)	10.8 (4.1)
Delay Detection Threshold (msec)	260.1 (110.8)	224.4 (97.3)	198.9 (78.9)
Synchronisation # of taps	27.6 (0.8)	28.1 (0.7)	28.0 (0.3)
Synchronisation vector length (z-transformed)	1.64 (0.4)	1.56 (0.4)	1.80 (0.3)
Synchronisation angle (msec)	-141.8 (63)	159.7 (67)	-86.6 (55)
Continuation # of taps	28.9 (2.2)	29.4 (1.6)	29.0 (1.1)
Continuation tempo drift (msec/sec)	-0.70 (1.7)	-0.81 (1.2)	-0.32 (0.9)
Continuation residual variability (% CV)	5.7 (3.4)	5.1 (1.5)	4.6 (1.4)
Continuation variability without de-trending (% CV)	5.9 (3.4)	5.3 (1.4)	4.7 (1.5)

Table 3: General auditory and motor test results for the three experimental groups. We report the mean (SD) unless otherwise indicated.

### 3.4 Discussion

We investigated whether auditory feedback can contribute to motor learning by assessing whether motor sequence regularity can be improved by keystroke-triggered sounds. In order to study learning transfer between sequences, participants were trained on one sequence (primary) on the first day and part of the second day, after which they briefly learned a second sequence (secondary).

#### 3.4.a Sequence learning and auditory feedback

A first set of key findings in this study relates to accuracy, speed and regularity of the sequence production. In the learning of the primary sequence, we found a reliable increase in the number of correctly tapped sequences per block (accuracy) as well as an increase in speed over the course of the first day. This trend continued on the second day when both quantities reached a performance plateau. In these quantities there was no difference in performance between the groups. Although the visual (yellow bar) metronome had indicated an inter-tap-interval of 300 msec prior to each trial, participants' actual

tapping speed was initially slower than this and throughout the learning phase approached the target speed. This means that perhaps we aimed too high and future studies could aim for a slightly slower tempo such as 350 msec or even 400 msec. On the other hand, it is a sign that our manipulation successfully provided a demanding test for our participants and therefore allowed them to learn.

The crucial variable of interest was the regularity of the tapping, defined as the standard deviation of the intervals between subsequent keystrokes. In order to correct for differences in tapping speed between participants, this quantity was expressed as a percentage of the average inter-keystroke-interval (for that participant and that block). We found that participants in the normal group (who received a woodblock tone at each keystroke) became increasingly regular in their tapping in the course of the first day. This improvement continued into the second day. The mute and jitter groups (who received no or jittered delayed feedback) showed no such improvement on the first day and only the mute group revealed a very mild improvement by the end of the second day (Figure 3 bottom).

In sum, all three groups showed improvements in accuracy and speed; but only the group who received immediate auditory feedback was able to improve in timing regularity.

The finding that accuracy and speed improve as participants perform the same sequence over and over again is widespread across the literature (Doyon, 2008; Nissen & Bullemer, 1987). Interestingly, we found no differences in performance between the groups upon the onset of auditory feedback. This contrasts with the finding of Conde et al. (2012) who found that participants benefited from auditory feedback already during the first block. The difference could be explained either by the fact that they tested musicians whereas our participants were nonmusicians, or alternatively the difference could be due to the fact that our task focused on tapping regularity instead of speed (which was the main dependent variable that Conde and colleagues focussed on with instructions asking for speeded RTs). Further, our finding squares with previous results indicating that the brain can use tone feedback for movement coordination timing (Ronsse et al., 2011) and our result adds that this is also true for sequence production. However, our study is the first to directly show that motor sequence learning is enhanced in a group that receives sound feedback relative

to a group who does not (the mute group). The only previous study who contrasted a sound with mute condition (Hoffmann et al., 2001) found greater learning for the sound group, but in their case, performance improvement could have been due to purely perceptual learning (Remillard, 2003). In our study, perceptual learning cannot explain the motor learning because the sequence to be learned was explicit and appeared on the screen throughout the experiment. Further, we add an important piece of evidence, which is the absence of improvement in the jitter group, which reveals that it is not the sound in itself but its timing relative to the movement that is necessary for learning. We will return to this latter point later on.

A question that needs to be briefly addressed at this point is why our mute group showed relatively little motor learning in timing regularity. The mute condition was designed to align with previous studies investigating motor sequence learning without augmented auditory feedback. As in previous SRT studies without tones (Conde et al., 2012; Nissen & Bullemer, 1987; Song, Howard, & Howard, 2008) the mute group *did* improve in terms of accuracy and speed, but our study adds to this observation that they did not improve in timing regularity. Most motor learning studies have restricted their attention to the parameters of accuracy and reaction time without considering regularity. It is interesting to note that the mute condition differs from other motor learning studies because we carefully eliminated all auditory feedback. Buttons always make sounds, taps are always heard, and most studies do not mask such auditory feedback because it does not yield a confound. The elimination of all auditory feedback in our mute group has therefore set it apart from previous motor learning study. As a matter of fact, in the context, we were surprised how difficult it was to dampen the keystroke noise and make it clearly inaudible (such as by supplying white noise over isolating headphones). This suggests that auditory feedback may not only be sufficient for motor learning (as we set out to show) but it may even be *necessary* for certain aspects of motor control (such as timing regularity). Targeted studies would be required to directly test this *tantalising idea*.

### 3.4.b Sequence-specific vs. unspecific learning

We investigated how much of what was learned on one sequence transferred to another sequence. The secondary sequence had the same formal properties as the primary sequence (it was the primary sequence for others) but was unknown to participants. If all learning is sequence-specific, we expected to observe performance values to return to the baseline level of the first block of the primary sequence. This was the case for sequence production accuracy in all but the normal group, whose participants were able to tap more correct sequences in the first block of the secondary sequence than for the primary sequence. Since each sequence was primary and secondary equally often this effect cannot be explained by properties of the sequences themselves. In terms of speed, the expected pattern was observed: performance in the first secondary block slowed down to levels similar to (albeit slightly faster than) the first primary block. This pattern was identical for the three groups. As for regularity, only the normal group showed an improvement between the first block of the primary and secondary sequence, clearly indicating sequence-unspecific learning. In sum, the normal group performed better on the secondary sequence in terms of accuracy, speed and regularity, whereas the other groups only improved in speed.

Why was only the group receiving sound able to learn sequence-unspecific accuracy and timing? First, this finding underlines the importance of testing sequence-unspecific learning as it may differ between experimental conditions. In the present experiment, participants' performance on the secondary sequence revealed that the normal group had improved sequence-unspecific motor control whereas the mute or jitter groups had not. In particular, sensory consequences of actions appear to contribute to transfer of skill between sequences, but only if the sensory consequences are time-locked with the movements. In our case, the transfer sequences were formally similar (except for the scale sequence, which we will discuss below) to the learned sequences. A further question is how transfer of learning depends on similarities between the learned and the transfer sequence. In our case, secondary and primary sequences were generated using the same structural constraints (see materials). Future studies could investigate how transfer differs between sequences that share more or less structural features.

### 3.4.c Auditory feedback deprivation

We investigated the effect of auditory feedback deprivation by adding a block after both the learning of the primary and secondary sequence in which the groups received no more keystroke-triggered sound anymore. This implied a condition change for the normal and jitter feedback groups, whereas the situation remained the same for the mute group. The result was that accuracy dropped in both the normal and jitter group (Figure 4). As for speed, the jitter group tapped the sequence faster in the case of both the primary and secondary sequence mute block. This finding is in line with the idea that they were obstructed by the auditory feedback and therefore improved their performance in its absence. On the other hand, given the *decrease* in accuracy, an alternative explanation is that they traded off accuracy for speed. In terms of tapping regularity, there were no clear within-group differences between the mute block and its preceding block. Instead, the normal group outperformed the jitter group on the secondary sequence without showing an effect of auditory feedback deprivation. Together with the finding that the normal feedback group alone was able to improve movement regularity, this suggests that the learning had become sufficiently stable as to no longer require the feedback during the second day. This finding is reminiscent of pianists whose timing performance was only mildly affected or not at all when they played on a mute piano (Gates et al., 1974; Repp, 1999b).

Both mute blocks occurred during the second half of the second day of learning. The finding suggests that at this stage of learning auditory feedback may be used to maintain accuracy but not regularity. For example, the feedback might serve as a signal that the keystroke has been successfully registered. Participants had received an alternative signal in the form of the fading of the corresponding digit on screen (see methods) but this was deliberately much less salient. Previously, the guidance hypothesis stated that when participants learn with feedback, their performance might degrade when this feedback is absent (Salmoni et al., 1984). The jitter group's improvement in speed might be a sign of what we would call a counter-guidance effect. We propose that once feedback is recognised by the motor system to be unreliable or otherwise counterproductive (as in the case of our jitter group) performance may improve when the feedback is absent. The idea that (constantly) delayed feedback is detrimental to performance is reported in (Cunningham,

Billock, & Tsou, 2001) who had participants learn play a video game where they needed to manoeuvre a plane through a field of obstacles using the computer mouse. They compared a condition of immediate feedback (35ms, the smallest interval possible given the hardware) with delayed feedback (235ms) and found that the latter condition impaired performance. But surprisingly, the group that had been practicing with delay feedback was impaired when feedback suddenly came immediately during the post test (a negative after-effect). In the present study, we did not test how the jitter group would perform under immediate feedback. This could be tested in a future study to answer the question whether noisy feedback could give rise to guidance effects. Further studies could be performed to investigate such effects by supplying feedback signals with increasing degree of distortion. The counter-guidance theory would predict that the improvement in performance in the absence of feedback (relative to performance with feedback) correlates positively with the degree of distortion of the feedback.

#### **3.4.d Scale playing transfer task**

The fourth key investigation of this study was performance in the tapping of a simple scale sequence (1234321) which was the same for all participants and measured before and after their learning of the two training sequences. The scale was tapped in one block paced by a preceding visual metronome and in a following block at maximal speed. Crucially, the scale tapping happened in the absence of keystroke-triggered auditory feedback; instead, all participants heard white noise that covered residual keystroke sounds. In this way, all groups were in the same condition and any tapping performance differences are attributable to transfer from the sequence learning and their feedback condition during this learning phase.

As for tapping accuracy in the paced scale blocks, only the jitter group deteriorated and the normal and mute group maintained the same performance level. The finding might reflect a speed-accuracy trade-off, with participants increasingly focusing on speed and timing regularity at the expense of sequence tapping accuracy. For the paced scale, similar to the development of sequence tapping speed, all groups tapped slower than the indicated 300 msec IOI on the first day but approached it closely on the second day. In terms of variability, the jitter and mute group tended to deteriorate in tapping regularity whereas the

normal group maintained the same level of regularity. As a result, there were group differences on the second day with the normal group outperforming the other two groups.

The result for the speeded blocks (in which participants tapped the scale as fast as possible) is clear: speed improves at the expense of accuracy and variability. This is likely due to a speed-accuracy trade-off. It reveals that participants respected our request to tap the sequence at maximal speed. No differences are apparent between the groups.

In sum, there is a strong tendency for an improvement in speed at the expense of accuracy and regularity, albeit to a lesser extent in the normal (accuracy and regularity) and mute (accuracy) groups. Whether these developments are due to the sequence learning in between the blocks, or perhaps alternatively due to fatigue remains open. For example, at the end of the two-day experiments participants may have been fatigued and therefore prefer speed over accuracy (Healy, Kole, Buck-Gengle, & Bourne, 2004). We have not tested a control group that was not trained in any sequence between the two days (as this was beyond the scope of this article).

### **3.4.e Does the sound need to be time-locked to the movement?**

The improvement in performance of the normal group indicates that auditory feedback benefits the learning of timing regularity in motor sequence production. We further show that it is not simply the fact that a tone is present that benefits motor performance. Previous accounts have argued that sensory feedback signals function as an aid to represent and select particular movements (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Shin, Proctor, & Capaldi, 2010). This reasoning was then used to explain benefit of sound on sequence learning (Hoffmann et al., 2001). Our findings provide evidence against such an explanation for two reasons. First, such explanation cannot explain why a delay in the sound as applied in the case of our jitter group would be unlikely to influence the sound's capacity to perform this role. This argument is strengthened by our finding that the range of jitter (10-190 msec) was just below the threshold for delay detection, making it unlikely that the delay influenced the creation of sound-motor associations. Participants in the jitter group did not mention noticing anything unusual about the sound in their debriefing questionnaire. Secondly, in our study all four keystrokes mapped to the same

sound. Therefore, the sounds could not be used to select individuated finger movements in the sense proposed by Hoffmann et al., 2001. In sum, our study provides evidence against action selection theories as an explanation for the role of auditory feedback in motor learning. Rather, we suggest that it is the time-lock between the auditory feedback and the motor action that is a crucial ingredient. This is what would be expected if the motor system uses the auditory feedback to update its motor program (Shadmehr & Krakauer, 2008). A limitation of our study is that we did not test a group in which the sound was delayed by a constant amount. The performance of this group would allow to distinguish two possible hypotheses: firstly, that the sound needs to be time-locked to the movement (in which case the constant delayed group would perform as well as our normal group), or secondly, that the sound needs to temporally coincide with the movement (in which case the constant delayed group would perform less well than the normal group).

#### **3.4.f Anisochrony, delay detection and synchronisation-continuation tapping tasks**

We found that the three groups did not differ in auditory temporal prediction precision (as measured by the anisochrony task) after the learning. The groups were also not different in their capacity to detect an asynchrony between their keystroke and a resulting tone. Furthermore, the three groups were matched in other features such as musical training (see Table 2). However, we found that the normal group was able to tap closer to the beat in the synchronisation-continuation tapping paradigm. Our interpretation for this latter effect is that the normal group had managed to learn the temporal association between their keystroke and the sound and was therefore able to make it occur closer in time to the metronome click. However, we also found that the groups did not differ in delay detection threshold. If our interpretation is correct, the normal group should have had a lower threshold in that task.

A limitation of these tests is that our participants performed them only after the learning part of the experiment on the second day. The reason we omitted these during our pre-testing (on day one) was that we wanted to avoid priming participants to pay attention to manipulations in the sound. As a result, however, our finding that the participants in the normal group tap closer to the beat the sensorimotor synchronisation task should be

treated with caution. It is possible that due to a sampling problem the populations were different in this respect at the outset. We feel that this is not likely because the synchronisation performance (tap-click distance) correlated with performance on the last three blocks of the sequence tapping experiment.

### 3.4.g Limitations and outlook

Our data provide evidence that auditory feedback can and does play a critical role in motor sequence learning. In our study, we found that immediate, time-locked auditory feedback improves motor learning relative to no auditory feedback or temporally jittered feedback, at least for learning timing regularity.

Our study has some limitations. We have not performed a retention task on the third day or after a week. In this way, the question remains open how long-lived the motor improvements here observed really are. Furthermore, although we aimed to keep participants' tapping speed constant across the experiment, we found that participants initially tapped slower than indicated. We feel that it is unlikely that this fact influenced our results, since we calculated the coefficient of variation which controlled for speed differences. A future study could use a slightly slower pacing signal to ensure that participants are able to achieve the intended speed from the first block onwards.

## 3.5 References

- Aschersleben, G., Gehrke, J., & Prinz, W. (2001). Tapping with peripheral nerve block. a role for tactile feedback in the timing of movements. *Experimental brain research*, 136(3), 331–339.
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods*, 37(3), 379–384.
- Bernier, P.-M., Chua, R., Bard, C., & Franks, I. M. (2006). Updating of an internal model without proprioception: a deafferentation study. *Neuroreport*, 17(13), 1421–1425. doi:10.1097/01.wnr.0000233096.13032.34
- Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal Of Experimental Psychology. Learning Memory And Cognition*, 16(1), 17–30. doi:10.1037/0278-7393.16.1.17
- Conde, V., Altenmüller, E., Villringer, A., & Ragert, P. (2012). Task-irrelevant auditory feedback facilitates motor performance in musicians. *Frontiers in Auditory Cognitive Neuroscience*, 3, 146. doi:10.3389/fpsyg.2012.00146
- Cunningham, D. W., Billock, V. A., & Tsou, B. H. (2001). Sensorimotor adaptation to violations of temporal contiguity. *Psychological science*, 12(6), 532–535.
- Deroost, N., & Soetens, E. (2006). Perceptual or motor learning in SRT tasks with complex sequence structures.

*Psychological Research*, 70(2), 88–102. doi:10.1007/s00426-004-0196-3

- Doyon, J. (2008). Motor sequence learning and movement disorders. *Current Opinion in Neurology*, 24(4), 478–483. doi:10.1097/WCO.0b013e328304b6a3
- Drewing, K., Stenneken, P., Cole, J., Prinz, W., & Aschersleben, G. (2004). Timing of bimanual movements and deafferentation: implications for the role of sensory movement effects. *Experimental brain research*, 158(1), 50–57. doi:10.1007/s00221-004-1870-9
- Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: influence of the tempo and musical backgrounds of listeners. *Brain and Cognition*, 58(1), 133–147. doi:10.1016/j.bandc.2004.09.014
- Finney, S. A. (1997). Auditory Feedback and Musical Keyboard Performance. *Music Perception: An Interdisciplinary Journal*, 15(2), 153–174.
- Finney, S. A., & Palmer, C. (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect. *Memory & Cognition*, 31(1), 51–64.
- Fisher, N. I. (1995). *Statistical Analysis of Circular Data*. Cambridge University Press.
- Gates, A., & Bradshaw, J. L. (1974). Effects of auditory feedback on a musical performance task. *Perception & Psychophysics*, 16(1), 105–109. doi:10.3758/BF03203260
- Gates, A., Bradshaw, J. L., & Nettleton, N. C. (1974). Effect of different delayed auditory feedback intervals on a music performance task. *Perception & Psychophysics*, 15(1), 21–25. doi:10.3758/BF03205822
- Ghez, C., Gordon, J., Ghilardi, M. F., & Sainburg, R. (1995). Contributions of Vision and Proprioception to Accuracy in Limb Movements, 19–36.
- Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes–no task. *The Journal of the Acoustical Society of America*, 93(4), 2096. doi:10.1121/1.406696
- Gu, X., & Green, D. M. (1994). Further studies of a maximum-likelihood yes–no procedure. *The Journal of the Acoustical Society of America*, 96(1), 93. doi:10.1121/1.410378
- Healy, A. F., Kole, J. A., Buck-Gengle, C. J., & Bourne, L. E., Jr. (2004). Effects of prolonged work on data entry speed and accuracy. *Journal of experimental psychology. Applied*, 10(3), 188–199. doi:10.1037/1076-898X.10.3.188
- Helmuth, L. L., & Ivry, R. B. (1996). When two hands are better than one: reduced timing variability during bimanual movements. *Journal of Experimental Psychology. Human Perception and Performance*, 22(2), 278–293.
- Hoffmann, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(2), 470–482. doi:10.1037//0278-7393.27.2.470
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *The Behavioral and brain sciences*, 24(5), 849–878; discussion 878–937.
- Hove, M. J., & Keller, P. E. (2010). Spatiotemporal Relations and Movement Trajectories in Visuomotor Synchronization. *Music Perception: An Interdisciplinary Journal*, 28(1), 15–26.
- Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology. Human Perception and Performance*, 36(6), 1525–1534. doi:10.1037/a0019059
- Hyde, K. L., & Peretz, I. (2004). Brains That Are Out of Tune but in Time. *Psychological Science*, 15(5), 356–360. doi:10.1111/j.0956-7976.2004.00683.x
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6), 718–727.
- Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correlational analysis. *Acta Psychologica*, 60(2-3), 173–191.
- Keller, P. E., & Koch, I. (2006). The planning and execution of short auditory sequences. *Psychonomic Bulletin & Review*, 13(4), 711–716.

- Koch, I., & Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction-time tasks. *Psychological Research*, 63(1), 22–35.
- Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244–247. doi:10.1038/nature02169
- Kunde, W., Koch, I., & Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *The Quarterly journal of experimental psychology. A, Human experimental psychology*, 57(1), 87–106. doi:10.1080/02724980343000143
- Kuriyama, K., Stickgold, R., & Walker, M. P. (2004). Sleep-dependent learning and motor-skill complexity. *Learning & Memory (Cold Spring Harbor, N.Y.)*, 11(6), 705–713. doi:10.1101/lm.76304
- Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical Plasticity Induced by Short-Term Unimodal and Multimodal Musical Training. *The Journal of Neuroscience*, 28(39), 9632–9639. doi:10.1523/JNEUROSCI.2254-08.2008
- Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon Symposium* (pp. 112–146). New York: Wiley.
- Münste, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews. Neuroscience*, 3(6), 473–478. doi:10.1038/nrn843
- Nezafat, R., Shadmehr, R., & Holcomb, H. H. (2001). Long-term adaptation to dynamics of reaching movements: a PET study. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, 140(1), 66–76.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1–32. doi:10.1016/0010-0285(87)90002-8
- Oscari, F., Secoli, R., Avanzini, F., Rosati, G., & Reinkensmeyer, D. J. (2012). Substituting auditory for visual feedback to adapt to altered dynamic and kinematic environments during reaching. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 221(1), 33–41. doi:10.1007/s00221-012-3144-2
- Pfordresher, P. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology. Human Perception and Performance*, 29(5), 949–964. doi:10.1037/0096-1523.29.5.949
- Pfordresher, P., & Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research*, 66(1), 71–79.
- Povel, D.-J., & Collard, R. (1982). Structural factors in patterned finger tapping. *Acta Psychologica*, 52(1–2), 107–123. doi:10.1016/0001-6918(82)90029-4
- Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of experimental psychology. Learning, memory, and cognition*, 29(4), 581–597.
- Repp, B. H. (1999a). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Perception & Psychophysics*, 61(3), 529–548. doi:10.3758/BF03211971
- Repp, B. H. (1999b). Effects of auditory feedback deprivation on expressive piano performance. *Music Perception*, 16(4), 409–438.
- Ronsse, R., Puttemans, V., Coxon, J. P., Goble, D. J., Wagemans, J., Wenderoth, N., & Swinnen, S. P. (2011). Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. *Cerebral cortex (New York, N.Y.: 1991)*, 21(6), 1283–1294. doi:10.1093/cercor/bhq209
- Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychological bulletin*, 95(3), 355–386.
- Schaffert, N., Mattes, K., & Effenberg, A. O. (2011). An investigation of online acoustic information for elite rowers in on-water training conditions. *JOURNAL OF HUMAN SPORT AND EXERCISE - University of Alicante*, 6(2), 392–405. doi:10.4100/jhse.2011.62.20
- Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, 185(3), 359–381. doi:10.1007/s00221-008-1280-5
- Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience*, 33(1), 89–108.

doi:10.1146/annurev-neuro-060909-153135

- Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological bulletin*, *136*(6), 943–974. doi:10.1037/a0020541
- Song, S., Howard, J., & Howard, D. (2008). Perceptual sequence learning in a serial reaction time task. *Experimental Brain Research*, *189*(2), 145–158. doi:10.1007/s00221-008-1411-z
- Stöcker, C., Sebald, A., & Hoffmann, J. (2003). The influence of response--effect compatibility in a serial reaction time task. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *56*(4), 685–703. doi:10.1080/02724980244000585
- Stuart, A., Frazier, C. L., Kalinowski, J., & Vos, P. W. (2008). The Effect of Frequency Altered Feedback on Stuttering Duration and Type. *Journal of Speech, Language, and Hearing Research*, *51*(4), 889–897. doi:10.1044/1092-4388(2008/065)
- Stuart, A., Kalinowski, J., Rastatter, M. P., & Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America*, *111*(5 Pt 1), 2237–2241.
- Tillmann, B., Stevens, C., & Keller, P. E. (2011). Learning of timing patterns and the development of temporal expectations. *Psychological Research*, *75*(3), 243–258. doi:10.1007/s00426-010-0302-7
- Ullen, F., & Bengtsson, S. L. (2003). Independent Processing of the Temporal and Ordinal Structure of Movement Sequences. *Journal of Neurophysiology*, *90*(6), 3725–3735. doi:10.1152/jn.00458.2003
- Waldstein, R. S. (1990). Effects of postlingual deafness on speech production: implications for the role of auditory feedback. *The Journal of the Acoustical Society of America*, *88*(5), 2099–2114.
- Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, *14*(1), 5–12. doi:10.3758/BF03198607
- Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature reviews. Neuroscience*, *12*(12), 739–751. doi:10.1038/nrn3112
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882. doi:10.1126/science.7569931
- Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in cognitive sciences*, *2*(9), 338–347.
- Wong, J. D., Kistemaker, D. A., Chin, A., & Gribble, P. L. (2012). Can proprioceptive training improve motor learning? *Journal of neurophysiology*, *108*(12), 3313–3321. doi:10.1152/jn.00122.2012
- Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin*, *60*(3), 213–232. doi:10.1037/h0044155

### 3.6 Acknowledgments

This work was supported by the EBRAMUS, European Brain and Music Ph.D. Grant to FTV (ITN MC FP7, GA 238157). The team "Auditory cognition and psychoacoustics" is part of the LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-60). We are indebted to Alexia Ferréol for invaluable assistance in running a subset of participants.

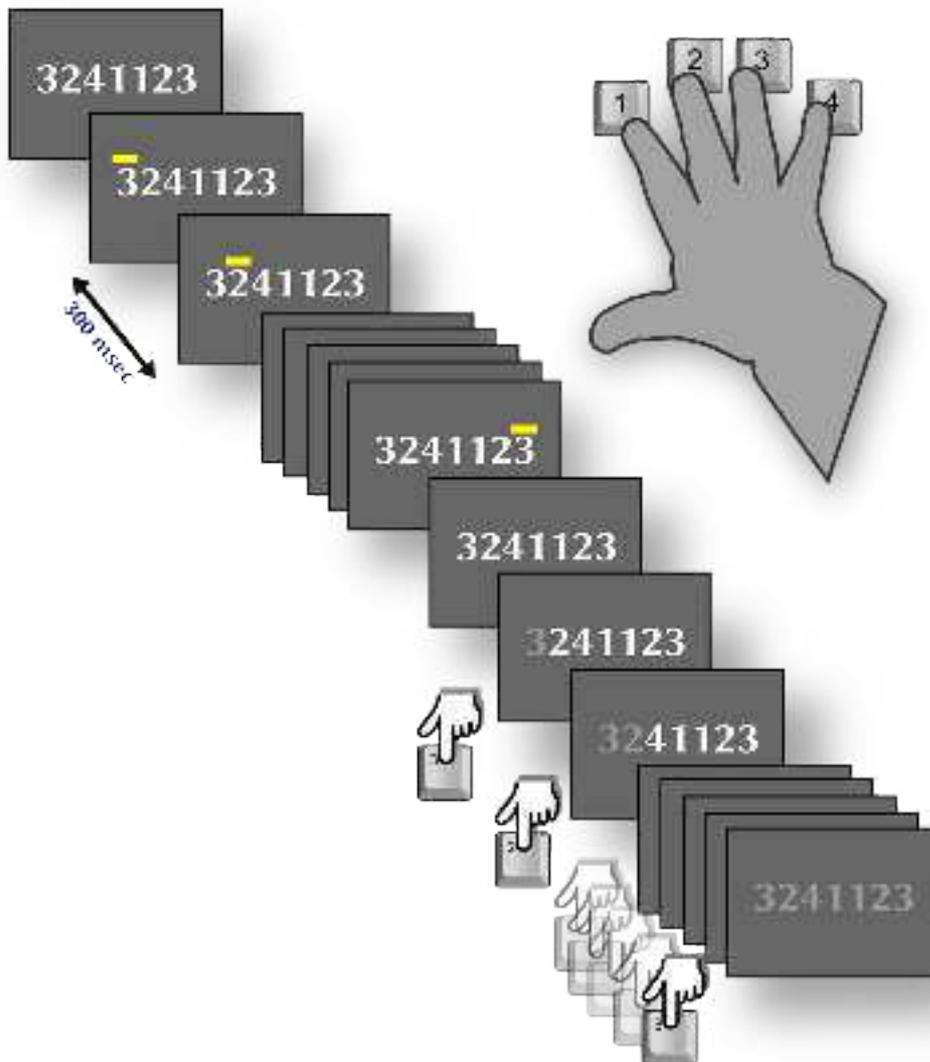


Figure 1: Experimental procedure on a single trial. A yellow overline first indicated the items of the to-be-produced sequence one by one at the intended speed (300 msec IOI). The participant then tapped the sequence on the four response buttons. After each keystroke, the item in the sequence was greyed out on the screen.

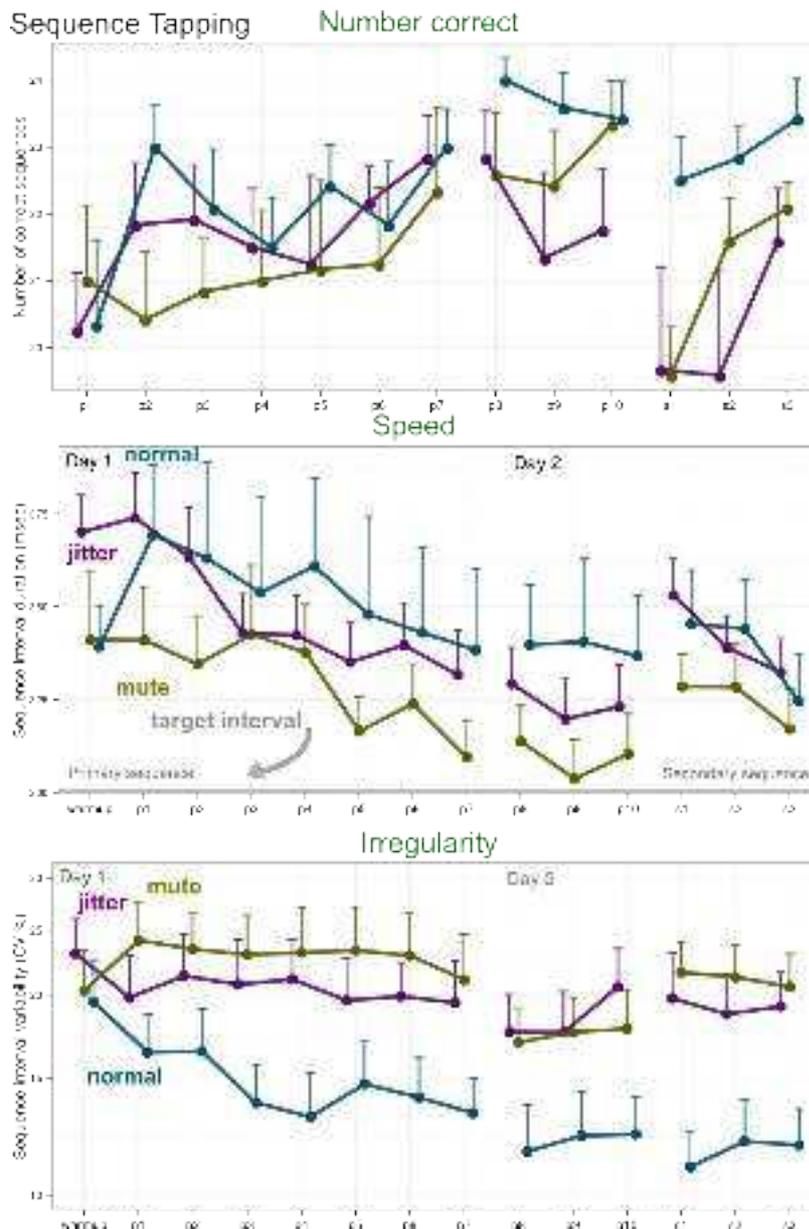


Figure 2: Speed and irregularity of the timing of the sequences trained during the two days. The top plot indicates speed as the average within-sequence interval duration (in msec). The bottom plot indicates the variability of the within-sequence inter-keystroke-intervals, expressed as a percentage of the average duration and presented on a logarithmic scale. The x-axis indicates the blocks, arranged chronologically. Here p1 to p10 indicate the 10 training blocks of the primary sequence and s1 to s3 indicate the training blocks of the secondary sequence.

Sounds on time - Learning in musical novices

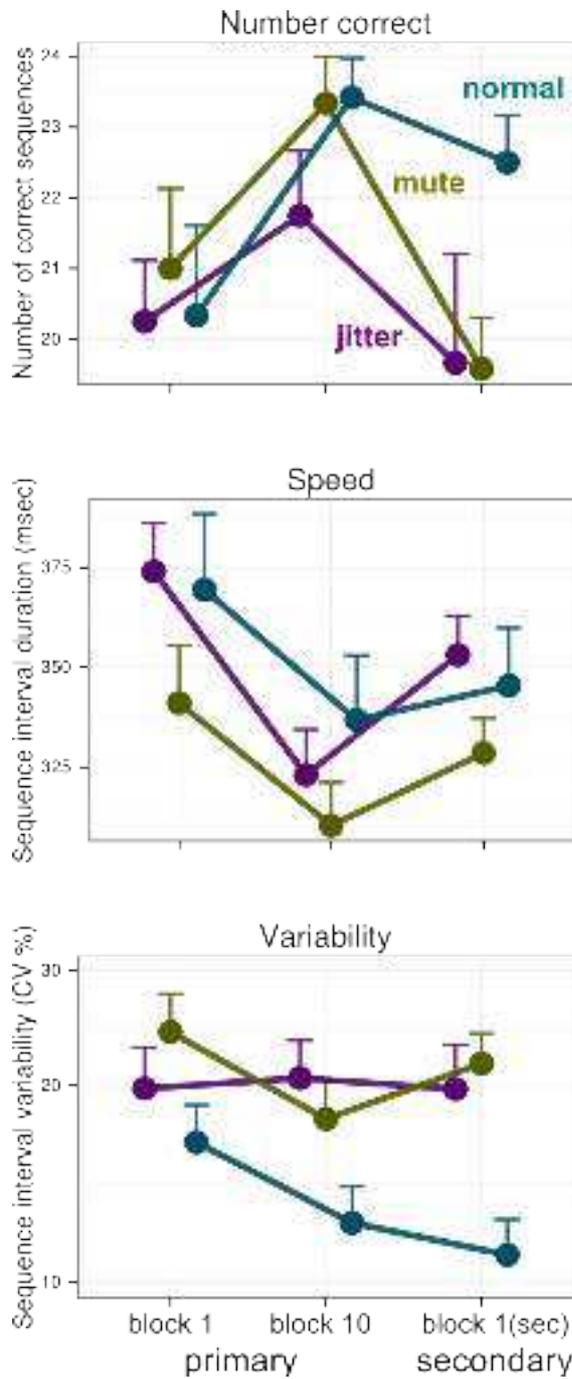


Figure 3: Investigating the effect of sequence switching. Significance levels: + $p < .1$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$  (significant levels in green indicate overall, group-independent) effects.

Sounds on time - Learning in musical novices

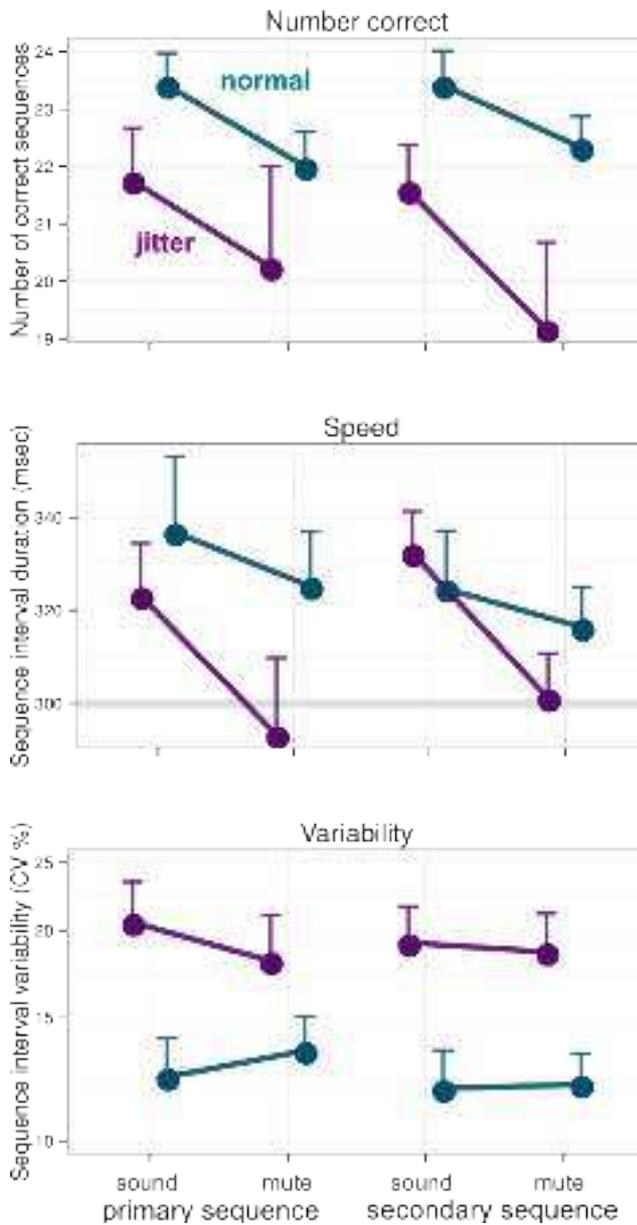


Figure 4: Investigating the tapping performance between the muted block (labelled mute) and the preceding non-muted (labelled by-group) block for the primary (left) and secondary (right) sequence. We report number of correctly produced sequences (top), tapping speed (middle) and tapping irregularity (bottom). The x-axis indicates the blocks in chronological order (omitting several intermittent blocks for clarity of presentation). Significance levels: + $p < .1$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Sounds on time - Learning in musical novices

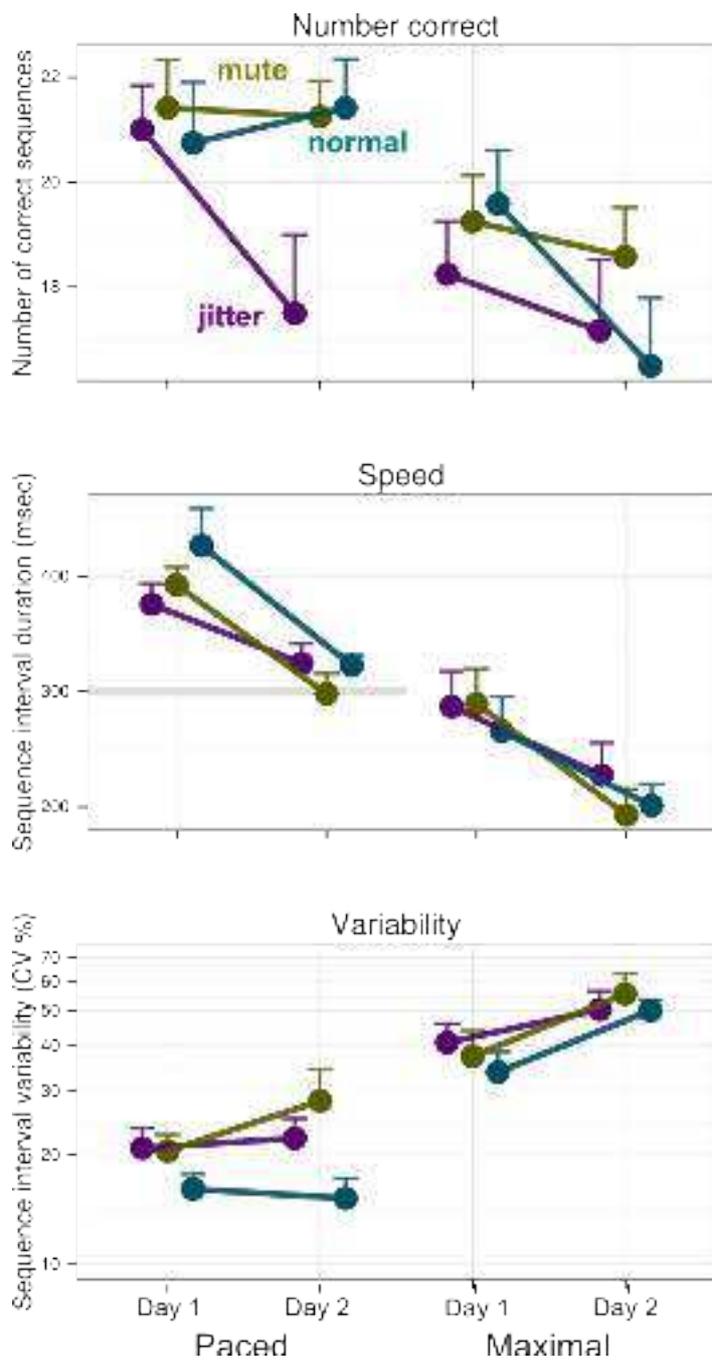


Figure 5: Tapping performance on the paced (300 msec IOI) and maximal (speeded) scale (1234321) sequence. Significance levels: + $p < .1$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

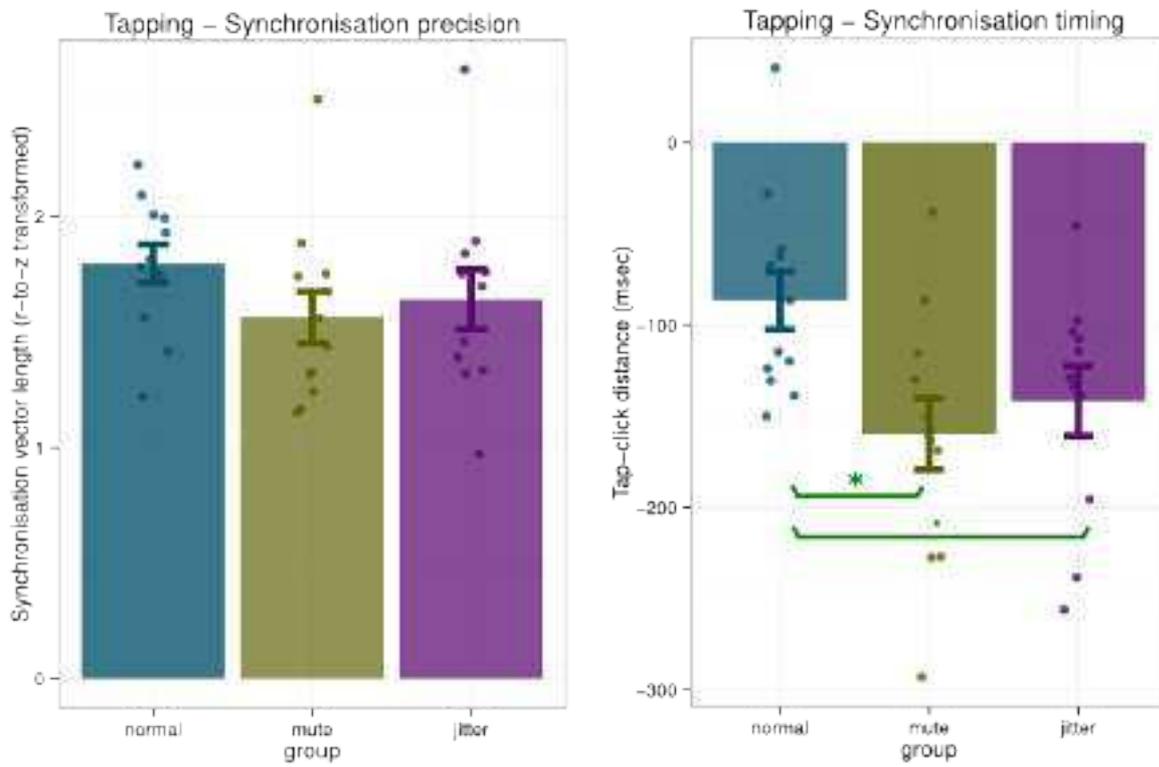


Figure 6: Synchronisation performance of the three groups. The left-hand plot indicates the synchronisation precision (a measure of the extent of time-lock between participant's tap and the metronome click). The right-hand graph indicates the temporal distance between tap and click, with negative values indicating that the tap preceded the click. For clarity of presentation, we have converted the synchronisation angle (in radians) back into a temporal distance (in msec). Indicated are means and standard errors. The main finding is that the three groups synchronise equally well but the normal group taps closer to the beat. Significance levels: + $p < .1$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

## **II.4. Discussion**

---

The present part of the thesis has investigated the role of auditory feedback in learning to tap a sequence regularly. The key findings are that we unequivocally implicate auditory feedback in motor learning.

The importance of these findings for the thesis as a whole will be discussed in the next and final chapter of the thesis.

### **III. Re-learning in stroke patients**

## III.1. Introduction

---

This part of the thesis will investigate the potential of music, and in particular auditory feedback, in the rehabilitation of stroke patients. We will briefly summarise the (vast) literature on motor stroke rehabilitation. The hypothesis that forms the basis for study III.2 in this thesis is that auditory feedback may supplant proprioceptive feedback in stroke patients undergoing music-supported stroke rehabilitation (Altenmüller, Marco-Pallares, Münte, & Schneider, 2009). Therefore, we will provide details about proprioception and its role in motor rehabilitation. Finally, we will present the aims and hypotheses for the studies presented in this part of the thesis.

### 1.1 Stroke: pathophysiology and rehabilitation

Stroke (or cerebrovascular incident) is a loss of or damage to brain tissue, typically as the result of a malfunction in the cerebral blood supply. For example, if a blood vessel bursts, parts of the brain are flooded with blood (haemorrhage) which causes tissue damage. Similarly, if a blood vessel is blocked (for example, by a clot) the tissue's oxygen and nutrient supply is cut off (ischaemia). Stroke is a major cause of disability worldwide (Dimyan & Cohen, 2011).

#### 1.1.a Causes of motor impairments

A majority of stroke patients suffer from motor deficits following their stroke. These can vary from paralysis of limbs to moderate or mild motor coordination deficits. Hemiparesis is the medical condition in which one side of the body is weakened. In the case of a stroke, this is the side of the body opposite to the hemisphere of the brain that is damaged by the stroke. Several interpretations have been advanced about what causes these deficits. For the sake of brevity, we will discuss a selection that are relevant to the

subsequent discussion.

It is generally thought that corticospinal tract integrity determines whether or not a patient remains disabled after stroke (Stinear et al., 2007). The spinal tract contains projections mostly from the primary motor cortex, but also from secondary motor areas such as the (dorsal) premotor cortex or supplementary motor areas (He, Dum, & Strick, 1993). It has been suggested that stroke damage to the primary motor cortex may cause the brain to recruit intact secondary areas (peri-lesional areas) (to replace the faulty primary cortex (Krakauer, 2006). Indeed, patients who showed poor recovery exhibited increased activation in non-primary motor areas (Ward, Brown, Thompson, & Frackowiak, 2003). This was interpreted to indicate that patients with particularly extensive primary motor area damage were able to recruit secondary motor areas.

*Contralateral inhibition* is the idea that the non-affected (i.e. contra-lesional) hemisphere's primary and non-primary motor areas become hyperactive in the acute phase of the stroke. This hyperactivity may then obstruct the functioning of M1 in the lesioned hemisphere. Indeed, inhibitory rTMS to the contralesional M1 improves the finger dexterity of the affected finger (that is, the finger that is operated by the lesioned hemisphere) (Nowak et al., 2008). Similarly, Ward, Brown, Thompson, & Frackowiak (2004) found that patients with stronger motor impairment have more bilateral motor activation during fMRI. They interpret this as showing that the other hemisphere takes over from the lesioned one. They also suggest that this picture changes for more chronic stroke patients (>3 months post-trauma).

Another idea aiming to explain stroke motor impairments is that the motor system fails to compensate for limb weakness following stroke (Mercier, Bertrand, & Bourbonnais, 2004). The authors hypothesise that stroke-induced hemiparesis symptoms are caused by limb weakness, combined with a failure of the motor system to update the internal models to compensate for this. They measure the maximum voluntary force (MVF) of patients and later ask them to produce some lower (submaximal) force with both sides of the body. They find (in line with previous research) that the paretic limb is systematically producing about 70% less force, although the patients were instructed to produce the same force with both sides. Further, since the force is submaximal, the lesioned hand has the capacity to produce force matching that of the healthy hand. This suggests that the motor system has not been

able to recalibrate the internal representations to reflect the motor loss. This limb force inequality (and lack of awareness thereof) has been suggested cause coordination difficulties (Lodha, Coombes, & Cauraugh, 2012).

In addition to purely motor deficits, stroke has been associated with a wide variety of sensory abnormalities. For example, many patients exhibit extinction, which is the phenomenon that patients fail to see a sensory stimulus if it is presented in both visual fields although they can detect it easily when presented to only one visual field (Urner et al., 2013). Additionally, patients often suffer from neglect (the inability to direct attention to objects in one visual field). Deficits are often accompanied by lack of awareness of these deficits on the part of the patients, and sometimes surprising degrees of denial (Vossel et al., 2012).

### **1.1.b Mechanisms underlying neural repair after stroke**

A certain amount of functional recovery after stroke happens spontaneously and those who exhibit more such spontaneous recovery are typically the ones having a more positive long-term outcome (Cramer, 2008). The amount of motor recovery is heterogeneous: some patients recover almost completely, whereas some remain impaired for the rest of their lives. Smania et al. (2007) show that whether the patient can extend their (paretic) finger is an important indicator of how much motor skill they will recover. Similarly, the Barthel index (a measure of the patient's independence in various daily tasks such as getting dressed or going to the bathroom) one week after stroke explains about 56% of the variance in functional outcome at 6 months post-stroke (Kwakkel, Kollen, van der Grond, & Prevo, 2003).

In general, there are two strategies in motor rehabilitation after stroke (Ward & Cohen, 2004). Firstly, patients are encouraged to learn to adapt to the existing impairments and secondly, they are offered therapies aimed at reducing the impairment itself. Since the neural tissue affected by the stroke is in principle damaged permanently, functional recovery relies on neuronal reorganisation, of which recruitment of peri-lesional areas is an example.

A number of therapeutic interventions are aimed at reducing the impairment. For example, reducing the somatosensory input from the intact (unaffected) hand improves

functional recovery of the affected hand (Floel et al., 2004; Werhahn, Mortensen, Van Boven, Zeuner, & Cohen, 2002). Similarly, increasing the somatosensory input from the affected hand improves its functioning (Conforto, Kaelin-Lang, & Cohen, 2002). Down-regulating the lesional primary motor cortex has beneficial effects (Siebner et al., 2000), probably due to reduced contra-lesional inhibition. Other approaches use pharmacological substances to accentuate effects of motor training (Bütefisch et al., 2002).

Arguably the gold standard of post-stroke therapeutic interventions is constraint-induced movement therapy (CIMT) (Shi, Tian, Yang, & Zhao, 2011; Taub, 2012) (although some meta-studies fail to find reliable effects; see (Corbetta, Sirtori, Moja, & Gatti, 2010; Langhorne, Coupar, & Pollock, 2009)). The idea behind this intervention is that patients will prefer to use their unaffected limb to perform daily activities. However, not using the limb will only deteriorate motor control (so-called learned non-use). Therefore, in CIMT the patient's *unaffected* limb is immobilised, forcing the patient to use the affected limb in activities of daily living. There is a consensus in the literature that CIMT is an effective tool for rehabilitation, at least for arm function but perhaps not for hand function (Langhorne et al., 2009).

Other researchers argue that stroke rehabilitation should be seen as a form of motor learning (Krakauer, 2006). In this way, it is argued that findings from the motor learning literature should be translated into clinical application. For example, stroke rehabilitation interventions typically focus on training patients in the movements of daily activity that they can no longer perform well. Motor learning studies, however, suggest that cortical changes in M1 may occur only through the learning of novel skills, and not just with repetitive use of old ones (Plautz, Milliken, & Nudo, 2000).

## 1.2 Proprioception and motor (re)learning

Previous studies have suggested a relation between proprioceptive deficit and the potential for auditory information to benefit patients in the framework of music-supported therapy (Altenmüller et al., 2009), a therapy in which patients learn to play musical instruments. Therefore, prior to introducing our studies, we will define proprioception and what role it might play in post-stroke motor rehabilitation.

### 1.2.a What is proprioception?

Proprioception is awareness of body position in time and space, through receptors in muscles, tendons, ligaments and fascia (Sherwood, 1997). Simply put, proprioception is what enables the body to know where its limbs are. The term was coined by Sherrington (1907). However, it has subsequently been used to mean a number of very different things: “position discrimination, weight discrimination, force discrimination, thickness discrimination, detection and discrimination of imposed movements, detection of direction of movement, and discrimination of movement patterns.” (Carey et al., 1996, p.1271). Authors often use the term proprioception interchangeably with tactile or somatosensory sensation.

Some researchers make a distinction between “static” position sense, which is knowledge about the relative positioning of limbs, and kinesthesia, which is the awareness of their (relative) movement (Johnson, Babis, Soutanis, & Soucacos, 2008). Other researchers distinguish between the joint position sense and proprioception (Hirayama, Fukutake, & Kawamura, 1999). Joint position sense (JPS) involves knowledge of the angle of each joint, which constitutes a raw position information. JPS will therefore only tell the brain the relative positions of the limbs. In order to know where the limb is in space, limb-to-limb-coordinate angles need to be translated into body-central coordinates then external coordinates. Furthermore, for that operation, it is necessary to take into account additional information such as limb length. That is, you need some map of your skeleton, a kind of *body image* (Gandevia, Refshauge, & Collins, 2002). Such body schema information is known to be affected by manipulations such as, for example, local anaesthesia which increases perceived finger size by 70% (Gandevia & Phegan, 1999). Therefore proprioceptive functioning is similarly affected by these manipulations.

### 1.2.b Proprioception tests

#### i. Principles of proprioception tests

Proprioception tests are divided into *analytical* and *functional* tasks. An analytical task measures simple proprioceptive acuity. A functional task investigates how well a person can use this proprioceptive information in a particular context. As an example, in a

typical analytical task, a participant's limb is moved according in a certain trajectory (invisible to the participant). Then the participant is asked to reconstruct this movement (the capacity measured here is also called repositioning sense). Alternatively, the limb is passively and invisibly moved by the experimenter to a particular location, and the participant is asked to point to the exact position, for example behind a screen. In a typical functional task, the participant stands on a surface, and needs to compensate for the varying tilt of the surface, again with eyes closed.

The analytical tasks can be divided into active and passive, according to whether the limb that the participant needs to locate is moved by the experimenter (passive) or by himself (active). For example, it is a well-known finding that a movement that is produced actively is reconstructed more accurately (Paillard & Brouchon, 1968; Winter, Allen, & Proske, 2005).

## ii. Main-stream proprioception tests

The test principles above are implemented in various ways (Carey et al., 1996).

In the joint-position sense (JPS) test, the experimenter moves the participants' limb, and asks her to match the position with the opposite limb (Dover & Powers, 2003).

In the *Thumb Localisation Test* (TLT) (Hirayama et al., 1999), the participant moves the patient's arm ("fixed limb") to a new location, and asks her to pinch the thumb of that limb with the opposite thumb ("reaching limb"). The findings for 221 stroke patients revealed TLT deficits when the limb contralateral to the cerebral lesion or ipsilateral to the peripheral nerve lesion (i.e. the affected limb) was the fixed limb, but not when the affected limb was the reaching limb. This indicates that the patients revealed deficits in localising their affected limbs, but not their unaffected limbs.

A proprioception test is furthermore included in the "sense of position" part of the Fugl-Meyer test battery (Fugl-Meyer, Jääskö, Leyman, Olsson, & Steglind, 1975). The experimenter moves a patient's limb passively (thumb, wrist, elbow and shoulder) and asks her to reproduce it. A variant on this is proposed by Carey et al. (1996) in which the experimenter moves a patient's wrist, for example, below a table. The patient is asked to point with their other hand on top of the table in real-time where they think their wrist. Errors are measured in degrees and therefore fairly reliable. Vidoni and Boyd (2009) use an

elegant test where the participant places her two hands on horizontal levers that are nearly frictionless. The experimenter moves one of the hands in a random fashion and asks the patient to copy the movement with the other hand whilst keeping her eyes closed.

### **1.2.c Alterations in proprioception**

Like most of our senses, proprioceptive functioning can be altered. In a particularly amusing manipulation, known as the *Pinocchio illusion*, one touches one's nose with closed eyes and at the same time a vibration is applied to the biceps tendon (i.e. the locus the biceps muscle attaches to the bone). The vibration makes the muscle spindles signal to the brain that the arm is moving away. At the same time one still receives the tactile information of touching one's nose. The way the brain makes sense of this information by assuming that one's nose has become longer (Lackner, 1988).

It is believed that cold temperature reduces proprioception. This would be a side-effect of *cryotherapy*, in which a limb is cooled down, for example after injury. Indeed, it seems that throwing-movement joint motion reproduction path deviates more after ice application (Wassinger, Myers, Gatti, Conley, & Lephart, 2007). However, a cold water immersion study found no effect on limb repositioning sense (Costello & Donnelly, 2011). Also, proprioception is believed to be decreased by nociception (pain perception), suggesting that they use similar pathways to the brain, which lead clinicians to train proprioception in pain therapy. Fatigue is argued to decrease proprioception. Indeed, participants are often subject to a fatigue protocol to simulate loss of proprioception (Thedon, Mandrick, Foissac, Mottet, & Perrey, 2011). After thalamic lesions (for example due to stroke), contralateral proprioceptive deficits are reported, sometimes surprisingly selectively, but reportedly with large variability (Gutrecht, Zamani, & Pandya, 1992; Sacco, Bello, Traub, & Brust, 1987). Finally, certain skin stimulations are thought to improve proprioception, for example the use of tape or bandages. These are widely used in athletics as a measure to boost the effects of training (Simoneau, Degner, Kramper, & Kittleson, 1997).

### **1.2.d Proprioception and motor learning**

Empirical evidence links proprioception to the motor system. Indeed, motor

learning appears to induce changes in limb position sense (Ostry, Darainy, Mattar, Wong, & Gribble, 2010; Vahdat, Darainy, Milner, & Ostry, 2011). Vice versa, and a proprioceptive training paradigm helps enhance motor performance (Rosenkranz & Rothwell, 2012). This enhancement is stronger if the training focused on passive instead of active proprioception (Wong, Kistemaker, Chin, & Gribble, 2012). These findings led researchers to propose that motor learning causes a recalibration of proprioception (Cressman & Henriques, 2009). This suggests that proprioception and motor learning are tightly linked in healthy participants.

However, Bernier et al. (2006) show that proprioception is not necessary for motor learning. A deafferented patient was able to learn a visual rotation in a reaching task. This suggests that visual feedback alone is sufficient to update an internal model, at least in the case of a lesioned system in which proprioception is not available.

### **1.2.e Proprioception after stroke and its role in rehabilitation**

Connell et al. (2008) report that 34-64% of early stroke patients show proprioceptive deficits. Six months after their stroke, patients show significant improvements in the upper extremity. In another study, 60% of stroke survivors were found to have sensory deficits (Schabrun & Hillier, 2009).

In the clinical conventional wisdom, proprioceptive deficits at stroke onset are a negative predictor of functional outcome. Sensory impairments in general have been identified as a negative predictor on outcome of rehabilitation after stroke (Carey, 1995). On the other hand, many properly controlled studies find no effect of sensory impairments on motor outcome. For example, Rand et al., (1999) compared a group of rehabilitation patients exhibiting only motor deficits with a group that exhibited both motor and proprioceptive deficits (the latter being measured according to the thumb-localisation-test). They followed this cohort over six weeks and found an overall improvement, no differences between the groups arise. Vidoni and Boyd (2009) show that stroke patients' learning of motor sequences correlates with their performance on a proprioception (position matching) task. The motor learning they measured was a continuous form of the serial reaction time task. In this task, patients were tracking a curve with their hand. In this curve, a particular segment was recurrent and it alternated with random segments. Proprioceptive acuity

correlated with the learning that was specific to the repeating segment.

In sum, some evidence suggests that proprioceptive functioning plays a role in motor rehabilitation. This led researchers to propose proprioceptive training for patients, hoping that improvements in proprioception would lead to improvements in motor functioning. Schabrun and Hillier (2009) review sensory (not just proprioception) retraining methods. They concluded that there is some support for the effectiveness of passive proprioceptive training, in which nerves are stimulated electrically, but not much support for active retraining, which uses sensory exercises.

### **1.3 Music-Supported Therapy: working hypotheses**

#### **1.3.a Music interventions in clinical settings**

Music has been used in a therapeutic intervention setting to address dementia (Vink, Birks, Bruinsma, & Scholten, 2003), mechanically ventilated patients (Bradt, Dileo, & Grocke, 2010) and acquired brain injury in general (Bradt, Magee, Dileo, Wheeler, & McGilloway, 2010). Conclusions are often discouraging, prevailing criticisms being that studies are poorly done in terms of research methods, and the randomly controlled studies that remain, taken together, are inconclusive.

#### **1.3.b Music-supported therapy**

More recently, a novel paradigm dubbed music-supported therapy was introduced as rehabilitation tool for stroke patients (Schneider, Schönle, Altenmüller, & Münte, 2007). Encouraging results were that patients improved their motor functioning after three weeks of active piano and drum training. For example, patients improved in finger tapping, block manipulation tasks and other clinical tasks meant to simulate everyday living situations. The therapeutic benefit of music-supported therapy was stronger than for Constraint-Induced training<sup>2</sup> (Schneider, Münte, Rodriguez-Fornells, Sailer, & Altenmüller,

---

2 However, the finding that the patients receiving CIMT training did not improve is puzzling. CIMT is the gold standard in motor rehabilitation and is widely found to improve motor functioning (Liepert, Bauder, Miltner, Taub, & Weiller, 2000). A possible reason is that patients were not randomly assigned to the

2010).

Motor improvements were accompanied by changes in inter-electrode coherence as measured by scalp EEG, suggesting that the neural oscillations from various parts of the patients' brain became more coherent (Altenmüller et al., 2009)<sup>3</sup>. However, since the authors did not localise the sources responsible for the oscillations it is impossible to establish whether the cortical changes were related to interactions between auditory and motor areas, or reflected a general cortical reorganisation. A more anatomically precise measure was reported in a single-case study of a patient undergoing music-supported therapy (Rojo et al., 2011). Brain scans of the patient revealed neural activation along the motor strip upon listening to the musical passages that he had learned to play during therapy. In parallel, improvements in motor performance were observed.

### **1.3.c Probing the mechanisms of music-supported therapy**

The question we ask in this part of the thesis is why music-supported therapy improves patients' motor functioning. What distinguishes music-supported therapy from other interventions? We argue that it is important to understand the reasons for its effectiveness, because such understanding could lead to improvements in the therapy and its application.

Researchers suggested that a crucial ingredient to music-supported therapy was sensorimotor integration (see part I of this thesis). But how exactly would sensorimotor integration contribute to motor rehabilitation? When one learns to play music, the brain learns to connect movements with sounds (see section I.5.4.e). That is, the mere occurrence of a sound leads to motor activation in the passive listener. As reported above, patients also show this auditory-motor co-activation after therapy (Rojo et al., 2011). Could this activation of the motor cortex due to sounds be responsible for the improvements in motor functioning? We argue that this is unlikely. Firstly, the motor activation is specific to when the individual is listening to sounds: it dissipates after the sounds stop. Therefore a prediction would be that patients are able to move well while they play (because of the

---

music or CIMT group, but rather first the music group was recruited and trained and then the CIMT group was recruited and trained.

3 It is important to note at this point that although these various studies in music-supported therapy exist, many are based on largely the same patient population.

sounds) but return to worse performance after they stop. This is not what was found. Instead, patients improved their non-musical movements (such as finger tapping, a box-and-box task, and a fine finger movement task, see Schneider et al., 2007). Secondly, the motor activation occurs *after* the movement is completed, since it is caused by the sound, which, in turn, is caused by the keystroke. Therefore it is predicted that it comes too late: the movement has already been performed and therefore cannot benefit from the auditory-to-motor activation anymore.

What are the alternative explanations? One interpretation prompted by the current state of research was that auditory-motor integration is somehow impaired in stroke patient populations (Rodriguez-Fornells et al., 2012). As evidence for this, they show that chronic patients do not exhibit motor activation during passive music listening prior to therapy (see section I.5.4.e). The auditory-motor connections, they argue, are restored through the learning of auditory-motor associations during music-supported therapy. Somehow, these connections would then improve patients' motor functioning. We argue that there is no evidence to date that the lack of motor activation during passive music listening is related to motor impairment. A neuroimaging study could measure this motor activation in a series of patients and attempt to find a correlation, which, if found, would provide evidence for this claim. To date, no such evidence exists. Further, the account fails to explain (a) how exactly auditory-motor connections benefit movement, (b) why other therapies (e.g., constraint-induced movement therapy) can be effective in restoring motor functioning, since they do not involve sound.

A second hypothesis proposes that the sound is used as a feedback signal to replace missing proprioceptive information in stroke patients (Altenmüller et al., 2009). The authors present the following argumentation. First, stroke patients are said to exhibit proprioceptive deficits. Second, the authors argue proprioception is beneficial for motor rehabilitation (although in the sections above we have shown that only limited evidence exists for this claim). Third, the idea is that patients' brains use the sounds generated by their movements to detect errors and correct these on subsequent trials (see section I.2.1.b). At present, no direct evidence exists for this claim, and therefore study III.2 in this section aims to test this hypothesis.

Thirdly, it is possible that the motor improvements are mediated by mood

improvements due to music. Music may put people in a more positive mood and thus influence recovery. Indeed, improvement in mood due to antidepressants has been shown to improve motor recovery (Chollet et al., 2011). In order for patients to maximally benefit from music's effects on mood, music-supported therapy (Schneider et al., 2007) has patients learn to play familiar tunes (such as children's songs) since familiar music is shown to elicit stronger emotional responses (Pereira et al., 2011). Related to this is the idea that music, and in particular playing together, helps to create social interaction between individuals (Overy, 2012; Wiltermuth, 2012; Wiltermuth & Heath, 2009). In particular, study III.3 in this section aims to test whether pairs of patients playing music together in synchrony could boost their music-supported rehabilitation.

## 1.4 Aims and hypotheses

We hypothesised that if auditory feedback is critically involved in music-supported motor therapy, patients should be affected by temporal distortions of the sound. Further, the auditory feedback that is inherent in music might provide the potential to boost rehabilitation by inviting pairs of patients to synchronise their sound.

The aims of this part of the thesis are the following.

1. Establish whether auditory feedback is critically involved in music-supported motor stroke rehabilitation (**study III.2**). We manipulated the auditory feedback in a way that was below patients' perceptual threshold and investigated the time-course of their rehabilitation.
2. Investigate whether manipulations in feedback affect patients' basic auditory, motor and auditory-motor functioning (**study III.2**).
3. Explore how synchronising the auditory feedback between pairs of stroke patients influence the benefit they experience (**study III.3**). We instructed one group of patients to play two voices of a musical piece together in synchrony, whereas a second group played in turns. To control for the potential benefit of sharing the

Sounds on time - Re-learning in stroke patients

rehabilitation experience, we had all participants play in pairs.

Sounds on time - Re-learning in stroke patients

### **III.2. Random delay is more efficient than immediate auditory feedback in fine motor rehabilitation after stroke**

---

**Random delay is more efficient than immediate auditory feedback in fine motor rehabilitation after stroke**

van Vugt FT<sup>1,2</sup>, Kafczyk T<sup>1</sup>, Kuhn W<sup>1</sup>, Rollnik JD<sup>3</sup>, Tillmann B<sup>2</sup>, Altenmüller E<sup>1</sup>

<sup>1</sup> Institute of Music Physiology and Musicians' Medicine, University of Music, Drama and Media, Hannover, Germany

<sup>2</sup> Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, CNRS-UMR 5292, INSERM U1028, University Lyon-1, Lyon, France

<sup>3</sup> BDH-Klinik, Hessisch Oldendorf, Germany

### Abstract

Motor impairments are among the most common and most disabling results of stroke worldwide. Previous studies have revealed that learning to play the piano helps to improve motor functioning in standardised motor tasks. One hypothesis about the effectiveness of the piano training is that it provides the brain a time-locked auditory feedback (a musical tone) with each movement (keypress). In our study, we provide the first direct test of this rehabilitation mechanism. We hypothesised that using a keyboard emitting randomly delayed tones, which thus fails to contain this temporal information about the movement, should affect the time-course of rehabilitation. Thirty-four patients in early stroke rehabilitation with moderate motor impairment and no previous musical background learned to play the piano using simple finger exercises and familiar children's songs. The participants were assigned randomly to one of two groups: in the *normal* group, the piano played the tone immediately at keypress, in the *delay* group, the tone occurred 100-600ms after the keypress. To assess recovery, we performed standard clinical tests such as the nine-hole-pegboard test and finger tapping rate and regularity. Furthermore, we used the Profile of Mood States (POMS) to establish motivational changes. Finally, we assessed whether patients were aware of the delay by testing keystroke-sound asynchrony thresholds, auditory temporal thresholds and by addressing qualitative questionnaires. Surprisingly, patients in the *delay* group improved more in the nine-hole-pegboard test than patients in the *normal* group. The groups improved similarly in finger tapping speed but the delay group outperformed the normal group in tapping regularity during therapy sessions. Music therapy on a randomly delayed keyboard can improve motor recovery after stroke, without affecting mood. Patients were unaware of the delay between their keystrokes and the resulting sounds. Our study suggests that sensory information may play an important role in motor rehabilitation. Possible explanations and implications for sensorimotor integration are discussed.

## 2.1 Introduction

Motor impairments are among the most common and most disabling consequences of cerebral stroke damage in humans (Dimyan & Cohen, 2011). The majority of rehabilitation interventions have focused on (re)training the motor system. In light of the fact that motor systems work closely in concert with sensory systems, we asked whether sensory (and in particular auditory) systems might contribute to rehabilitation.

Patients undergoing conventional therapy are shown to benefit from additional musical training (Schneider, Schönle, Altenmüller, & Münte, 2007). In these studies, patients learn to play the piano and a drum set to practise fine- and gross motor skills, respectively. Relative to a control group that received constraint-induced movement therapy, the musical intervention group showed greater improvement in a variety of clinical measurements of fine motor control (Schneider, Münte, Rodriguez-Fornells, Sailer, & Altenmüller, 2010; Schneider et al., 2007). These behavioural improvements were accompanied by evidence of neuroplastic changes in cortico-cortical coherence and movement-related potentials (Altenmüller, Marco-Pallares, Münte, & Schneider, 2009), corticospinal integrity (Amengual et al., 2013), and motor cortex excitability (Grau-Sánchez et al., 2013). However, it remains unclear why music training was more efficient than existing non-musical therapies.

We identify three possible explanations for the benefit of music-supported rehabilitation. Firstly, a possible explanation is that the movements involved in piano and drum playing simply constitute efficient exercises (*movement-only explanation*) and that these alone suffice to explain the rehabilitation benefit of music-supported therapy. Put differently, music therapy is just finger exercises. In its strongest formulation, such a hypothesis would entail that the sounds could be omitted, they play at best a peripheral role in rehabilitation. Secondly (or possibly additionally), the improvement in motor control in the patients may be explained by participants' greater engagement (*motivation explanation*) (Thompson, Schellenberg, & Husain, 2001). Music training is, after all, novel for all participants (all studies so far focused on patients without musical experience). One should not underestimate the motivational value for patients who, for the first time since their stroke, learn something that they were not able to do before their stroke, instead of learning

to approximate skills that they could do much better previously. Thirdly, it is possible that the connection between movement and resulting sound (auditory feedback) plays a key role in rehabilitation (*auditory-motor explanation*).

One hypothesis for this third explanation is that music training allows patients to benefit from auditory feedback supplied by the sound (Altenmüller et al., 2009) in a way that was not possible in non-musical rehabilitation therapies. The rationale was twofold. First, EEG-experiments in healthy adults, who are piano novices demonstrated that piano training resulted in auditory–motor co-activation already after the first twenty minutes of piano lessons (Bangert & Altenmüller, 2003; Lahav, Saltzman, & Schlaug, 2007). Indeed, such a co-representation of sounds and movements has been demonstrated in expert musicians (e.g., Bangert et al., 2006; Haslinger et al., 2005; Haueisen & Knösche, 2001). The establishment of such an auditory-motor co-representation was therefore believed to improve motor output in patients suffering from motor impairments. The reasoning was as follows. For motor rehabilitation after stroke, proprioception, the body's limbs position sense (Sherrington, 1907), is an important, although not necessary, prerequisite (Bernier, Chua, Bard, & Franks, 2006). Indeed, motor learning has been associated with changes in proprioceptive acuity (Wong, Kistemaker, Chin, & Gribble, 2012; Wong et al., 2012) suggesting a link between proprioception and motor learning. Stroke patients often exhibit deficits in proprioceptive perceptual function (Connell, Lincoln, & Radford, 2008) and such deficits correlate negatively with the patient's rehabilitative outcome (Carey, Oke, & Matyas, 1996; Schabrun & Hillier, 2009; Vidoni & Boyd, 2009). Those patients who cannot use proprioceptive feedback or can use it but to a lesser degree might be able to benefit particularly from auditory feedback as a basis for their motor rehabilitation. The idea that auditory feedback would replace proprioceptive feedback is similar to studies that supplant one sensory modality for another, such as auditory for visual (Bach-y-Rita, Collins, Saunders, White, & Scadden, 1969). In healthy individuals, there is evidence that auditory feedback can be used in learning to adapt to force field disturbances in reaching movements (Oscari, Secoli, Avanzini, Rosati, & Reinkensmeyer, 2012) or in timing of out-of-phase bimanual coordination patterns (Ronsse et al., 2011). This suggests the brain can use auditory feedback in motor learning in a similar way to how it uses visual and proprioceptive feedback. Possibly, auditory feedback is even superior to other kinds of

feedback because humans seem to become less reliant on it (Heitger et al., 2012).

The three explanations introduced above are not mutually exclusive and the actual rehabilitation benefit might be a combination of several of these factors. In the present study, we decided to focus on the third, auditory-motor, explanation. Recent evidence for auditory-motor explanation of the therapeutic benefit in music-based stroke rehabilitation may come from neuroimaging data that reveals that patients exhibit motor activation when passively listening to music that they have been trained to perform (Rodriguez-Fornells et al., 2012; Rojo et al., 2011). Such motor co-activation is a common finding in professional pianists listening to piano sounds (Bangert et al., 2006; Baumann et al., 2007; Haueisen & Knösche, 2001) and appears after brief periods of music-like training in non-musicians (Bangert & Altenmüller, 2003; Lahav et al., 2007). The observation shows that the stroke patients have learned to associate movements with piano sounds in the course of music therapy. However, it is unclear whether this auditory-motor activation played a critical role in the therapy or whether it is a side-effect.

The present study investigates this question directly by manipulating the auditory-motor coupling between motor actions and sounds and observing its effect on the therapy outcome. The question arises what control group one should test. Initially, one might consider contrasting a group of patients receiving therapy on a silent (mute) piano with patients receiving therapy on a normal piano. However, patients will likely be strongly influenced by noticing that a piano is supposed to emit sounds. Furthermore, patients will not experience the same amount of motivation because they will not actually experience the result of their playing, that is, the music. Passive exposure to music has been shown to improve cognitive capacities in stroke patients (Särkämö et al., 2008). Therefore, we felt that a mute control condition would not do. A second alternative would be to use delayed feedback by programming the piano to give out each tone only after a certain delay. However, the human brain is shown to quickly adapt to sensorimotor delays by recalibrating sensory and motor synchrony, at least in the case of healthy individuals (Sugano, Keetels, & Vroomen, 2010). Therefore, a fixed delay would not amount to much manipulation in sensory feedback. We therefore opted to introduce a random delay between the keystroke and the sound, by choosing a random delay that was different for each keystroke. On a keyboard programmed in this way, a perfectly temporally regular tone

sequence would sound highly irregular.

In sum, to date no evidence exists that auditory feedback is actually used by these patients in music-supported therapy. The aim of the present study is to provide the critical test by degrading the auditory feedback in one group whilst keeping it intact in another and investigating the difference in rehabilitation progression between these groups. Answering this question will be crucial, not only for musical rehabilitation, but for exploring the rehabilitation potential of sensorimotor integration as well as helping us to understand the nature of the crosstalk between auditory and motor areas. Furthermore, contrary to previous studies (Altenmüller et al., 2009; Schneider et al., 2007), we decided to focus entirely on the piano playing part in musical rehabilitation so as to not confound differential effects of feedback in drum pads and the piano. A further aim of our study was to investigate motor skill changes in patients during the course of therapy. All previous studies in music-supported therapy have only administered motor tests before and after therapy and therefore crucial information is missing as to at what time in the therapy the movement changes occur and whether patients reach a plateau after several sessions or not.

## 2.2 Methods

### 2.2.a Patient population

Patients were recruited in one of the largest Northern German neurological rehabilitation hospital (BDH-Hospital, Hessisch-Oldendorf, Germany). They were undergoing conventional therapies in parallel to the therapy program we offered, described in detail below.

We aimed at obtaining a representative sample of patients from the hospital population. Consequently, we were not able to maintain high homogeneity of participant selection in terms of lesion localisation or time since stroke. However, our results are maximally relevant to clinical practice. Inclusion criteria were:

- light to moderate motor impairment in the upper extremity due to ischaemia or

haemorrhage;

- having residual movement capacity enabling the patient to move the arm and the index finger of the affected hand independently;
- age between 30 and 75 years;
- Barthel Index at least 50;
- right-handedness;
- at most 3 months post-stroke;
- being able to understand and agree with informed consent to participate.

Exclusion criteria were:

- previous musical training for more than 4 years;
- psychiatric problems (as assessed by standard clinical investigation);
- cognitive impairment or aphasia.

We assigned patients to one of two groups in a quasi-randomised design. Both groups received music therapy with exactly the same program. The only difference was that in the normal group, the keyboard emitted its sounds immediately upon keystroke, whereas in the delay group, the keyboard emitted sounds delayed by a random delay sampled from a flat distribution between 100 and 600 msec. This distribution of jittered delay was chosen so as to be as large as possible (in order to maximally disrupt the auditory-motor time-lock) whilst not being noticeable to the patient. It may seem surprising that a delay of 100-600 msec was not noticeable, but a follow-up questionnaire revealed that nearly all patients were unaware of the delay. The groups were slightly uneven because we initially included 3 patients whose stroke had happened too long ago (>200 days). Table 1 indicates the characteristics of our two groups. Unfortunately, all patients suffering from haemorrhage were allocated to one group. We have therefore calculated crucial tests both with and without these patients, suggesting that the stroke type does not seem to influence the effects reported hereafter.

	Delay group	Normal Group	Comparison
N. of patients	19	15	n/a
Gender (female/male)	9/10	6/9	Fisher exact p=.74
Affected hand (left/right)	9/10	6/9	Fisher exact p=.74
Post-stroke interval (days)	33.5 (18.3)	39.2 (28.7)	Mann-Whitney U, Z=-.07, p=.94
Stroke type (haemorrhage /ischemia)	0/19	5/10	Fisher exact p=.01
Barthel score (0..100)	52.9 (20.6)	70.6 (30.8)	Mann-Whitney U, Z=-1.54, p=.12
Education (years)	9.8 (2.1)	10.3 (1.3)	Mann-Whitney U, Z=-1.56, p=.12
Lesion site: (present/absent)			
- left frontal	0/19	1/14	Fisher exact p=.44
- left occipital	0/19	1/14	Fisher exact p=.44
- left parietal	2/17	3/12	Fisher exact p=.63
- left temporal	2/17	1/14	Fisher exact p=1.00
- left subcortical	11/8	6/9	Fisher exact p=.49
- right frontal	3/16	4/11	Fisher exact p=.67
- right occipital	0/19	0/15	Fisher exact p=1.00
- right parietal	2/17	2/13	Fisher exact p=1.00
- right temporal	0/19	1/14	Fisher exact p=.44
- right subcortical	7/12	4/11	Fisher exact p=.72
Participation in other therapies during the study (h)	37.8 (14.1)	44.4 (15.5)	Mann-Whitney U, Z=-1.49, p=.14

Table 1: Clinical and biographical characteristics of our patient groups. We report mean values (standard deviation) unless otherwise indicated.

Our study took place in two phases. In the first phase, we measured 14 patients (7 in the normal group, 7 in the delay group). In a second phase, 20 more patients (12 delay, 8 normal) were measured and we included additional tests (such as auditory and sensorimotor tasks). Therefore, data on these tasks is only available on a subset of our patient pool.

## 2.2.b Therapy

Patients received 10 sessions of half an hour of piano training. The sessions took place on different days over the course of three to four weeks. The day before the first session and the day after the last session was dedicated to the PRE and POST measurements, respectively, which are described in more details below.

The therapy program followed the same structure every day. In the beginning, patients played simple finger exercises such as a five-tone scale up and down and other, alternating patterns (see Supplementary Materials for a detailed description of the program). Then patients learned to play one song from a set of simple children's songs. We chose to work with songs that were familiar because those are shown to elicit more emotional responses (Pereira et al., 2011). If patients reached a sufficient level on one of the songs, they would be invited to learn additional songs from the set prepared by the therapist (see Supplementary Materials).

Patients played on a M-Key V2 MIDI controller keyboard that was chosen for its light touch. The keyboard was connected through the M-Audio Midisport Uno MIDI-to-USB converter to a Linux laptop. The laptop ran a custom made C program that recorded the MIDI events to a file and forwarded it to the software synthesiser Fluidsynth which was equipped with a Steinway sound font and generated the sounds. The program additionally changed the MIDI velocity value (loudness) to its maximum value. As a result, all sounds sounded maximally loud, which was done to keep very soft keystrokes from being heard. Furthermore, in the delay-group the sounds were delayed by a number of milliseconds sampled newly for every keystroke from a flat distribution between 100 and 600 msec. That is, all values between 100 and 600 msec, inclusively, were equally probable. In the normal-group the sounds were played immediately (0 msec delay). The sounds were then played through Creative Inspire T10 speakers (Creative Labs, Inc.) at a comfortable loudness level. The five keyboard keys used in the therapy were numbered. Songs and exercises were then written in a simplified musical notation as numbers in tabular form and presented visually to the patients as a memory aid (see supplementary materials for an example).

Patients played the piano with the hand of their affected extremity only. The

therapist stood next to the patient and supported his or her forearm when so required. The patient was always encouraged to make as much of the movements by him- or herself as possible. For those patients who were more severely affected, the therapist initially pointed to the fingers or moved them gently, encouraging the patient to make the movements unassisted on the next trial. Throughout therapy, the therapist role was one of supporting and the aim was always to allow the patient to function as independently as possible instead of dependent on the therapist.

### **2.2.c Nine-hole pegboard test**

The nine-hole pegboard test is a standard clinical test for fine motor control in which a patient needs to place 9 small sticks (pegs) in 9 holes and take them out again (Heller et al., 1987; Mathiowetz et al., 1985; Parker, Wade, & Langton Hewer, 1986). The patient is seated comfortably with his affected arm resting on the table. The 9HPT board is placed so that he or she can reach it easily, with the side with the peg container at the side of the tested arm. The instruction is to take the pegs one by one from the board and put them in the holes, and then take them out and put them back, again one by one. The experimenter holds a stopwatch that is started once the patient touches the first peg, and is stopped once the patient releases the last peg. This test is performed once at PRE and once at POST.

### **2.2.d Finger tapping measurements**

We investigated patient's finger tapping performance of the affected hand as a measure of fine motor control. Four different tapping conditions were measured: (1) paced thumb-to-index tapping, (2) unpaced thumb-to-index tapping (continuation), (3) index finger speed tapping, (4) middle finger speed tapping. The tests are described in detail in what follows. In all conditions, patients were seated comfortably at a table on which they rested their arm. In order to have a portable, flexible and yet maximally accurate measurement of finger tapping performance, we custom designed a measurement device. Finger motion was recorded by a triaxial accelerometer (ADXL 335) attached gently to the patient's index or middle finger tip (depending on the task). Tap contact was measured by a force sensitive resistor (FSR SEN09375), which is a small sheet whose electrical resistance

changes upon contact in a way that is related to the contact force. Both sensors were read out by an Arduino Duemilanove experimentation board running a custom C program to sample sensors at 3 kHz. The data was then transferred online to a Linux laptop running a custom Python program interfacing through USB and allowing the therapist to preview the data. We made the blueprints of the device setup as well as the custom programs used to capture the data available freely online for future research groups to use (<http://www.github.com/florisvanvugt/immotion>).

In paced thumb-to-index tapping, patients were instructed to tap as regularly as possible in time with a metronome (Korg TM-40 Digital Tuner Metronome) at 69 BPM (i.e. 1.15 Hz) during 60 seconds from the first tap (Calautti et al., 2006). The thumb-to-index tapping movement was argued to be more natural and a more reliable reflection of activities of daily living (Okuno, Yokoe, Akazawa, Abe, & Sakoda, 2006). The side of the hand (at the little finger) rests lightly on the table and the fingers are held in a relaxed posture (not at maximum flexion or extension). The index finger and thumb move to touch each other, and then move apart again, to a distance of about 5cm (but at least 2cm).

In unpaced thumb-to-index tapping, the metronome was switched off but patients were asked to continue tapping at the same speed and regularity for another 30 seconds. This is similar to the continuation phase in synchronisation-continuation tapping paradigms (Wing & Kristofferson, 1973). This measure was only performed in the second phase of our study and therefore we only have this data for a subset of patients (see Patient methods).

In index finger speed tapping, we measured the maximum tapping rate and variability during approximately 14 seconds (measured from the first finger tap). The patient was invited to choose where it is most comfortable to tap on the table. His or her elbow rests on the table and the hand lies palm down on the table. The fingers are held in a relaxed posture close to maximal extension but slightly bent so that the position can be sustained without muscular effort. No metronome is used. The patient is instructed to tap as fast and as regularly as possible, lifting his or her finger at least 2cm above the table on each cycle. The force sensor surface is placed on the table and the patient is instructed to tap on the same spot.

In middle finger speed tapping, the procedure is the same as with index finger speed tapping but switching to the middle finger.

The raw data files containing the force trace over time were preprocessed using a custom developed python script. The script first discarded the first and last .5 sec of data from the recordings and then converted the force sensor trace into Newtons using a prior calibration table. It then smoothed the signal using a 160-sample Bartlett window, which amounted to approximately 53 msec at our sample rate. The script then detected the tap onset landmarks (a sudden impact) when the force exceeded .05 Newton. Tap offset (a release of contact from the tapping measurement surface) was detected when the force trace dropped below .05 Newton again at least 40 msec after the last tap onset. Similarly, the next onset was restricted to occur at least 75 msec after the last tap offset (see supplementary materials for more details). All data files with their landmarks were then visually inspected and in a number of cases the .05 Newton onset/offset threshold was adjusted manually to compensate for the fact that some patients tapped too softly. We furthermore recorded the maximal tapping force between subsequent tap onset and tap offsets; the intervals between adjacent onsets, which will be referred to as the inter-tap-intervals (ITIs) in what follows; and the duration between the tap onset and tap offset (tap dwell phase duration). Next, we discarded the ITIs that were larger than 2000 msec since these marked pauses or interruptions in the patient's tapping behaviour (such as asking the experimenter whether to continue tapping) instead of the patient's motor control capacity. We also discarded ITIs shorter than 120 msec which were considered artefacts.

### **2.2.e Patient mood measurements**

We used the short form of the Profile of Mood States (POMS) with 35 items (Lorr, McNair, & Droppleman, 1971). The short form has 35 adjectives (items) of which the patient rates to what extent they are applicable to his or her mood over the last week, on a scale from 1 (not at all) to 5 (very strongly). The items load onto four category sub-scores: depression/anxiety, fatigue, vigor and hostility (Curran, Andrykowski, & Studts, 1995). We used the validated German translation (Bullinger, Heinisch, Ludwig, & Geier, 1990). This questionnaire was measured at PRE and POST. The experimenter read each of the items to the patients, who then responded verbally. The responses were filled out on a sheet and

were at a later stage entered into a table independently by two researchers. The tables were compared and discrepancies corrected by checking again the paper forms.

In order to obtain a quick estimate of the patient's mood throughout the therapy, we used a mood scale of faces (Andrews & Crandall, 1976; Kunin, 1955; McDowell, 2006). The patients were presented a list of faces that are abstract smileys, ranging from very happy to very sad (see Supplementary Materials). The patients were asked at the PRE measurement session which face best represents how they were feeling by pointing to the corresponding face. At the beginning of each therapy session, the patients were asked to point to the face corresponding to how they had felt since the previous session. At the end of each therapy session, too, the patients were asked again how they felt during the therapy session. Finally, during the POST measurement the patients were asked to select the face representing how they felt at present. The therapist wrote down the letter code corresponding to the chosen face without allowing the patients to see the letter in question.

### **2.2.f Auditory and sensorimotor tasks**

In order to establish basic auditory functioning and auditory-motor processing in patients, we included a threshold measure of detecting purely auditory timing perturbation (anisochrony detection) and a measure of sensitivity to delays between auditory and motor events (auditory-motor delay detection). These were only measured in the second phase and therefore we only have the data for a subgroup of our patients.

In the anisochrony detection task, patients were asked to judge whether a sequence of five notes, in which the fourth note was sometimes delayed, sounded regular or irregular. We established the corresponding psychophysical threshold. Patients were seated comfortably and received headphones (Shure SRH440), set to a comfortable sound level, which was kept constant across all participants and all conditions. On each trial, they heard a sequence of five isochronous sine wave tones of 100 msec (with linear ramp of 5 msec) at an IOI of 350 msec, where possibly the fourth is played with a certain delay but the fifth tone was always on time (Ehrlé & Samson, 2005; Hyde & Peretz, 2004). That is, when the tone was delayed by an amount  $d$ , the third interval was longer by  $d$  msec and the fourth interval was shorter by  $d$  msec. The patient's task was to report whether the sequence sounded regular (isochronous) or not. The experimenter registered the verbal answer

(regular or irregular) in the computer. A maximum-likelihood procedure (MLP) (Green, 1993; Gu & Green, 1994) implemented in Python was used to adaptively control the stimulus level and converge to patient's threshold. For each block, the algorithm outputs an estimate for the participant's threshold. The MLP algorithm briefly works as follows. A set of candidate psychometric curves are maintained in parallel and for each, the likelihood of the set of the participants' responses is calculated. The psychometric curve that makes the participant's responses maximally likely is used to determine the stimulus level (the delay between the keystroke and the sound) on the next trial. We used 200 candidate psychometric curves with midpoints linearly spread between 0 and 200ms delay and crossed by the five false alarm rates (0%, 10%, 20%, 30%, 40%). Hence, a total of 3000 candidate psychometric curves were used.

Participants first performed 4 trials (2 with no delay and 2 with a delay of 200ms) to make clear the difference between a regular and irregular sound sequence. The participant received accuracy feedback about his or her answers during these practice trials. Next, they performed a block of 10 trials, starting at 200ms delay but then using MLP to determine the stimulus levels of the following trials. If the procedure was clear, we continued with 3 experimental blocks of 36 trials. Each experimental block consisted of 36 trials containing 6 catch trials. Catch trials are trials on which the delay was always 0 msec (regardless of the delay that was suggested by the MLP algorithm). The function of catch trials is to prevent participants from always responding “irregular” (which would cause the MLP algorithm converge to a zero threshold). Catch trials were inserted randomly as follows: the first 12 trials contained 2 catch trials and the next 24 trials contained 4 catch trials.

In the delay detection task, we measured patients' sensitivity to delays between auditory (tone) and motor (keystroke) events. That is, we established from which delay onwards patients noticed that the tone came after the keystroke instead of immediately. This test was performed only during the POST measurement and not during the PRE measurement so as to not alert patients to a possible manipulation of feedback during the therapy. In the *active* variant of this test, the procedure was as follows. We tested the patients' affected hand. Patients were seated comfortably and instructed to keep their eyes closed (otherwise they would have been able to use the visual information from seeing their keystroke and compare it with the onset of the sound). They rested their wrist on a slightly

elevated surface, reaching their index finger down to a small USB keypad (Speedlink SL7430 SGY). The finger was curved but held in a relaxed, natural way and rested on the key without having to force it down (or holding it up) (see Supplementary Materials for details). To support it, the hand rested on a wooden box (16.5 cm x 12 cm x 3 cm). The instructions were to raise the finger several centimetres from the key and then press down in a swift manner. The keypad was connected through USB with a Linux Laptop that emitted a woodblock sound through the headphones after a certain delay (in msec). . The patients responded verbally whether or not they had the feeling that the tone was delayed, and this was entered in the computer by the experimenter. We had first of all 4 trials (2 with no delay and 2 with a delay of 1200ms) to make the difference clear. The patients received feedback about their answers during this block. Next, a block of 10 trials was presented, starting at a keystroke-tone delay of 1200ms but then using MLP to determine the next stimulus level of each trial. If the patients understood, we continued with 3 experimental blocks of 36 trials using the MLP algorithm, again starting at 1200ms. As parameters for the MLP procedure, we used 500 hypotheses with midpoints linearly spread between 0 and 1000ms delay and crossed by the five false alarm rates. Hence, a total of 2500 hypotheses was used. At each trial, the patient pressed the key when he or she wanted to, and heard a tone. This tone was either played at the same time of the keypress or delayed by a number of milliseconds (determined by the MLP algorithm)

In the *passive* variant of this task, the procedure was exactly the same, except that in this case the patients were not invited to make a voluntary movement to press the key. Instead, the therapist held the patients' (affected) finger gently at four distinct points and pressed it down (see Supplementary Materials). Again, patients kept their eyes closed. Our reasoning was that if their proprioception was impaired, patients would be particularly affected in this task, since no sensory information other than proprioception could be used (except perhaps tactile information, which was reduced through the gentleness of the therapist's touch).

## **2.2.g Intra-therapy measurements**

The keyboard keystrokes were recorded by the custom C program that captured the MIDI information (see above). These transcripts were saved and analysed offline for all

patients in the second phase (subgroups). Whenever the therapist played the keyboard to show a movement, he held down the highest or lowest key on the keyboard at the same time. We used this signal tone to later discard keystrokes occurring within that time window and analysed exclusively notes played by the patient. Our analysis focused on the scale playing movement (the movement 123454321 where the fingers are numbered from 1=thumb to 5=little finger for the right hand or inversely for the left hand).

### **2.2.h Qualitative interview**

At the POST measurement we included a qualitative questionnaire. The patient's responses were recorded, transcribed and then synthesised. The patient is asked about the following: (1) how he or she experienced the therapy; (2) whether he or she feels the therapy has helped to improve his or her movement capacities, daily functioning or mood; and (3) what could be improved about the therapy, including questions investigating whether patients in the delay group were sensitive to the manipulated piano sound. For details, see the Supplementary Materials.

### **2.2.i Statistical analyses**

We performed parametric ANOVA whenever the data quantity and distribution could reasonably be assumed to fulfil its assumptions. We detected deviations from sphericity using Mauchley's Test and whenever it was significant we applied the Greenhouse-Geisser correction. In those cases, we indicated significance as  $p_{GG}$  and omitted the uncorrected p-value for brevity. Follow-up comparisons were performed using Tukey's HSD contrasts. We report generalised effect sizes  $\eta_G^2$ .

### **2.2.j Ethics**

This study was performed in accordance with ethical guidelines proposed by the Medical University Hannover (MHH). The protocol was approved by the ethics board on 20.4.2011 (nr. 1056-2011).

## 2.3 Results

### 2.3.a Nine hole pegboard test

We recorded the time (in seconds) for the patients to complete the nine hole pegboard test. Due to the non-normality of the data, we log-transformed the score. Patients in the delay group completed the test during the PRE session in 4.24 (SD=0.83) log seconds and patients in the normal group in 3.96 (SD=0.66) log seconds. This difference was not significant [ $t(32.0)=1.12$ ,  $p=0.27$ ]. We then performed an ANOVA with the time to complete the test as dependent variable and time point (PRE/POST) and group (normal/delay) as factors. We found no main effect of group [ $F(1,32)=.67$ ,  $p=.42$ ] but a main effect of time point [ $F(1,32)=30.41$ ,  $p<.0001$ ,  $\eta^2=.02$ ], which indicated that both groups improved between PRE and POST. Crucially, there was an interaction between group and time point [ $F(1,32)=4.91$ ,  $p=.03$ ,  $\eta^2=.003$ ], revealing that the delay group improved more than the normal group (Figure 1). Comparisons for the two groups separately indicated that the delay group improved significantly [ $t(18)=-5.63$ ,  $p<.001$ ,  $d=1.29$ ] whereas the normal group showed only a trend to improvement [ $t(14)=2.00$ ,  $p=.07$ ,  $d=0.51$ ]. In non-transformed scores, this amounts to an average improvement of 31.2 (median 14, SD=53.6 sec) sec in the delay group and an improvement of 8.7 (median 3.8, SD=17.9) sec in the normal group. The effect persisted even when we discarded patients who were more than 1.5 SD away from the mean of the two pooled groups during the PRE measurement [ $F(1,28)=3.97$ ,  $p=.06$ ] (discarding 4 patients, 3 in the delay group). With a discarding threshold of 3 SD no patients were discarded. The effect furthermore remained mostly intact when we eliminated the haemorrhage patients (see introduction) [ $F(1,27)=3.54$ ,  $p=.07$ ,  $\eta^2=.12$ ].

### 2.3.b Finger tapping measurements

We had to eliminate one patient (normal group) because of an unsuccessful PRE measurement due to technical difficulties and another patient (delay group) because the recorded taps were so soft that they could not be reliably distinguished from noise.

*Speed tapping measures.* To establish the speed of tapping, we calculated the median ITI for each patient and each measurement time point. We chose not to use the mean as the

median is less susceptible to outliers (long ITIs) than the mean.

During the PRE measurements, patients tapped at an average median ITI of 540 (SD=312) msec with their index finger, corresponding to a rate of approximately 1.85 Hz. This tapping speed was not different between the two groups [ $t(21.5)=-.43$ ,  $p=.67$ ]. We continued to perform a repeated-measures ANOVA with median tapping ITI as dependent variable and factors group (delay/normal) and time point (PRE/POST). We found a main effect of session [ $F(1,30)=19.11$ ,  $p=.0001$ ,  $\eta^2=.17$ ], indicating that patients tapped faster at POST (M=333, SD=118 msec ITI, corresponding to a rate of 3.00 Hz) than at PRE. There was no main effect of group [ $F(1,30)=.09$ ,  $p=.77$ ] nor an interaction between group and session [ $F(1,30)=.37$ ,  $p=.54$ ].

How did the tapping speed develop during therapy? We had recorded index finger tapping with maximum speed and regularity before and after each of the ten therapy sessions. In order to determine whether tapping speed differed between the measurements before and after the session, we ran an ANOVA with factors session (pooling sessions 1 to 5 and 6 to 10) and time (before or after). There was a small but significant effect of time [ $F(1,30)=6.88$ ,  $p=.01$ ,  $\eta^2=.001$ ], indicating that patients' tapping was slightly faster immediately following the therapy session (M=343, SD=132 msec ITI) than immediately before (M=353, SD=136 msec ITI). There was no effect of group and no interaction of time with group [both  $F(1,30)<.62$ ,  $p>.43$ ]. This showed that the groups did not respond differently to the individual therapy sessions in terms of tapping speed.

In order to simplify presentation of changes across the ten therapy sessions, we have pooled the tapping data recorded before and after each session (Figure 2). We performed an ANOVA with session (from 1 to 10), group (normal/delay) as factors. There was a main effect of session [ $F(9,270)=10.68$ ,  $p_{GG}<.001$ ,  $\eta^2=.06$ ] indicating that tapping became faster in the course of therapy. Indeed, Bonferroni-Holm-corrected post-hoc testing revealed that both groups became faster between the PRE measurement and session 1 [ $F(1,30)=10.25$ ,  $p(\text{corr})=.04$ ,  $\eta^2=.03$ ] and between sessions 1 and 2 there was only a trend for improvement [ $F(1,30)=8.11$ ,  $p(\text{corr})=.07$ ,  $\eta^2=.01$ ]. None of the later pairwise comparisons were significant [all  $F(1,30)<3.81$ ,  $p(\text{corr})>.54$ ]. This suggests that the majority of the tapping speed improvement happened early in the therapy (Figure 2). There was no main effect of group [ $F(1,30)=.26$ ,  $p=.61$ ] nor any interaction between group and session [ $F(9,270)=0.99$ ,

$p=.44$ ], indicating that both groups were tapping equally fast during therapy. In sum, we found that index finger tapping speed increased most importantly in the beginning of the therapy; that is, within the first few therapy sessions.

As for the middle finger tapping, patients tapped slower than with the index finger at an ITI of 583 (SD=328) msec during the PRE measurement and this was not different between the two groups [ $t(28.0)=.01$ ,  $p=.99$ ]. The same repeated-measures ANOVA as above yielded a significant effect of time point [ $F(1,30)=26.10$ ,  $p<.001$ ,  $\eta_G^2=.17$ ] but no effect of group [ $F(1,30)=.04$ ,  $p=.83$ ] nor an interaction [ $F(1,30)=.13$ ,  $p=.72$ ].

In sum, both groups improved in their tapping speed in index and middle finger tapping and there were no differences between the groups. The vast majority of this improvement in tapping speed appeared to take place in the first 2 sessions with only mild improvements afterwards.

*Tapping variability.* Estimating tapping variability can easily be biased by outlier values, since it essentially reduces to estimating the variance of a population of tapping intervals. For this reason, we discarded ITIs that were more than 3 SD away from the mean ITI for that session (this discarding is known as Extreme Studentised Deviation outlier rejection). However, we verified other common discarding algorithms (such as the boxplot rule, or no discarding at all) and they essentially yielded similar results, suggesting that our results are not caused by a specific outlier rule. In order to correct for possible tapping speed differences, we calculated the coefficient of variation (CV) by dividing the standard deviation of the ITIs by the median of that session (since the median is more stable). Since these variability values showed a skewed distribution, we furthermore log-transformed these coefficient-of-variation in order to satisfy parametric analysis assumptions and obtain more normally distributed groups. For readability, we report the untransformed coefficient-of-variations in percent (of the inter-tap-interval). A higher CV value indicates more variable tapping, a lower CV indicates less variable, more regular tapping.

First, we compared the PRE and POST measurement sessions. For this, we calculated an ANOVA with within-participant factor session (PRE/POST) and between-participant factor group (normal, delay). For the index finger tapping regularity, we found a main effect of session (PRE/POST) [ $F(1,30)=10.90$ ,  $p=.002$ ,  $\eta_G^2=.13$ ], revealing that patients were tapping more regularly after therapy ( $M=8.8$ ,  $SD=1.8$  %CV ITI) than before ( $M=13.4$ ,  $SD=1.7$  %CV

ITI). There was no effect of group [ $F(1,30)=.24, p=.63$ ] nor an interaction between group and session [ $F(1,30)=.48, p=.49$ ]. As for the middle finger, contrary to the index finger we found no improvement in variability between PRE and POST [ $F(1,30)=1.28, p=.27$ ] nor an effect of group [ $F(1,30)=0.14, p=.71$ ] or interaction [ $F(1,30)=.03, p=.86$ ].

Next, we investigated how finger tapping variability changed in the course of therapy. Again, we first of all checked for differences in tapping regularity between the index finger tapping measure preceding each therapy session and the measurement following it. To this end, we computed an ANOVA with within-participant factor session (10 levels for the 10 therapy sessions), within-participant factor time point (before or after the therapy session) and between-participant factor group. We found no effect of time (before or after the therapy session) [ $F(1,30)=0.29, p=.59$ ]. Therefore, we proceeded to pool the two recordings made during each therapy sessions. This allowed us to perform interval discarding (Extreme Studentised Discarding (ESD), that is, discarding all intervals more than 3 SD away from the mean for that patient and session) on the pooled data, which improved the signal-to-noise ratio. We found a main effect of session (1 to 10) [ $F(9,270)=2.39, p_{GG}=.03, \eta_G^2=.04$ ] revealing that overall, the tapping became more regular (less variable) in the course of therapy. Crucially, we also found a main effect of group [ $F(1,30)=4.70, p=.04, \eta_G^2=.07$ ], indicating that the delay group was tapping more regularly ( $M=10.7, SD=1.5$  %CV ITI) than the normal group ( $M=14.4, SD=1.4$  %CV ITI).

As is clear in Figure 2, the groups gradually move apart during therapy. Then, quite surprisingly, the normal group improves dramatically in regularity between session 10 and the POST measurement. In order to test whether these differences were significant, we performed a post-hoc ANOVA with factors time point (10<sup>th</sup> therapy session, POST measurement) and dependent variable tapping regularity. This revealed that the normal group improved significantly in tapping regularity from the 10<sup>th</sup> to the POST measurement [ $F(1,14)=7.57, p=.02, \eta_G^2=.11$ ] whereas the delay group stayed at the same [ $F(1,16)=0.00, p=.96$ ]. This explains why we did not find an interaction in tapping regularity between the two groups between PRE and POST.

*Tapping force and dwell duration.* Next, we investigated the tapping force. We

calculated the maximum force during each tap contact and then the median of these forces for each measurement session. First, we computed an ANOVA with time point (PRE/POST) and group and found a main effect of session [ $F(1,30)=10.00$ ,  $p=.004$ ,  $\eta^2=.11$ ]. This indicated that patients' tapping force had *decreased* at POST ( $M=40.7$ ,  $SD=37.8$  Newton) relative to PRE ( $M=24.4$ ,  $SD=18.5$  Newton). A second ANOVA for the measurements during the ten therapy sessions revealed that the tapping force remained constant during therapy [ $F(9,270)=.86$ ,  $p_{GG}=.52$ ] and there was no effect of group [ $F(1,30)=.04$ ,  $p=.84$ ] nor an interaction between group and therapy session [ $F(9,270)=.99$ ,  $p_{GG}=.43$ ].

Secondly, we investigated the variability of the maximum force (as coefficient-of-variation). Here, one patient's recordings had to be discarded because they contained many tapping force outliers, probably due to tapping on the border of the force sensor. An ANOVA yielded no main effect of session (PRE/POST) [ $F(1,29)=1.09$ ,  $p=.31$ ] or group [ $F(1,19)=.00$ ,  $p=.94$ ] nor any interaction [ $F(1,29)=2.54$ ,  $p=.12$ ]. The force variability furthermore remained stable throughout therapy [ $F(9,261)=1.36$ ,  $p_{GG}=.24$ ] and there was no main effect of group [ $F(1,29)=0.19$ ,  $p=.66$ ] nor an interaction [ $F(9,261)=.72$ ,  $p_{GG}=.60$ ].

Finally, we investigated the duration of the contact phase (sometimes referred to as dwell phase) of the tap. That is, how long did patients hold down their finger in the movement? We calculated the duration of the dwell duration as a percentage of the inter-tap-interval (ITI), since it has previously been associated with motor precision (Hove & Keller, 2010). We found no difference between the PRE and POST measurements [ $F(1,30)=2.29$ ,  $p=.14$ ] nor an effect of group [ $F(1,30)=1.00$ ,  $p=.33$ ] nor an interaction [ $F(1,30)=0.16$ ,  $p=.69$ ]. There was variation in the dwell duration during therapy [ $F(9,270)=3.28$ ,  $p_{GG}=.008$ ,  $\eta^2=.02$ ], reflecting an increase in dwell duration from 44.7 (SD 9.6) %ITI in session 1 to 49.2 (SD 8.1) %ITI in session 10. There was no main effect of group [ $F(1,30)=1.39$ ,  $p=.25$ ] or interaction between group and session [ $F(9,270)=0.72$ ,  $p_{GG}=.61$ ].

*Paced tapping.* Now we turn to the paced (and unpaced continuation) tapping tasks. In this task, patients tapped with their index to thumb fingers. However, for one patient (normal group) the PRE measurement overall force level was too low for reliable identification of tap onsets and similarly for two patients (one in the normal group, one in the delay group) the POST measurement force levels were too low. Therefore, we discarded

these patients from the analysis. Patients often had to skip a metronome click and therefore there were less taps registered during the PRE measurement ( $M=45.9$ ,  $SD=14.2$ ) than at POST ( $M=51.5$ ,  $SD=8.8$ ) [ $F(1,28)=5.54$ ,  $p=.03$ ,  $\eta_G^2=.06$ ]. Because of these missing taps, calculating classical measures such as the variability of inter-tap-intervals is problematic. We therefore turned to circular statistics (Fisher, 1995). Since we had no recording of the metronome click times, we could only calculate the time-lock (synchronisation vector length) but not the average phase (synchronisation angle) of the taps. The vector length is a measure of the time-lock between the taps and the metronome clicks and its range is from 0 (no time-lock) to 1 (perfect time-lock). We transformed this quantity using Fisher's r-to-z transformation to obtain a quantity that satisfies the ANOVA assumptions for equal variance. We found no effect of time point (PRE/POST) [ $F(1,28)=2.37$ ,  $p=.14$ ] nor of group [ $F(1,28)=.08$ ,  $p=.77$ ] and no interaction effect [ $F(1,28)=.83$ ,  $p=.37$ ]. That is, patients' synchronisation capacities remained essentially the same during our study, although they were able to produce slightly more taps after the therapy.

A subset of our patients (those in the second phase: 11 patients in the delay group and 8 in the normal group) also performed a continuation phase of this tapping movement, aiming to continue the same tapping rate and regularity. We first discarded tapping intervals that were more than 3 SD away from the mean ITI for that recording session, amounting to 1.2% ( $SD=1.3\%$ ). The amount of discarded material did not differ between the groups [ $F(1,17)=1.87$ ,  $p=.19$ ] nor time points [ $F(1,17)=0.21$ ,  $p=.65$ ]. As is typical in tapping continuation phases, (healthy) participants speed up or slow down. Therefore, the standard deviation of the inter-tap-intervals is inflated because it does not take into account the temporal correlation of the tapping intervals. Therefore, we detrended the ITIs by fitting a straight line and investigating the residuals from this line. The average slope of this line was  $-1.09$  ( $SD=3.8$ ) msec ITI/sec, with the negative meaning a decrease of the intervals (and therefore a slight speeding up). Across participants, this trend was not significantly different from zero [ $t(17)=-1.2$ ,  $p=.24$ ]. The residual variability after regressing out the individual trends was  $8.9$  ( $SD=7.7$ ) %CV ITI and was not different between groups [ $F(1,16)=.72$ ,  $p=.41$ ] or time points [ $F(1,16)=.00$ ,  $p=.99$ ] and there was no interaction [ $F(1,16)=.43$ ,  $p=.52$ ].

### 2.3.c Patient mood measurements (profile of mood states)

The results for the Profile of Mood States are presented in Table 2. Patients in the two groups did not differ in their Profile of Mood States (POMS) scores at PRE for depression/anxiety [ $t(30.1)=-.95, p=.35$ ], fatigue [ $t(29.3)=-1.10, p=.28$ ], vigor [ $t(30.8)=.65, p=.52$ ] or hostility [ $t(31.4)=-.07, p=.95$ ]. We then proceeded to run one repeated-measures ANOVA for each of the four POMS subscales with factors group (normal/delay) and time point (PRE/POST). We found a main effect of time point for depression/anxiety [ $F(1,32)=7.68, p=.009, \eta^2=.009$ ] and for fatigue [ $F(1,32)=8.49, p=.006, \eta^2=.05$ ]. This indicated that both groups were less depressed/anxious and had less fatigue after the therapy relative to before the therapy. There were no improvements in vigour [ $F(1,32)=1.29, p=.27$ ] or hostility [ $F(1,32)=.06, p=.81$ ]. There were no main effects of group [all  $F(1,32)<.91, p>.34$ ] nor any interaction effects between group and time point [all  $F(1,32)<1.62, p>.21$ ], indicating that mood was the same in both groups and did not develop differently in the course of the therapy.

	Depression	Fatigue	Hostility	Vigor
<i>Normal group</i>				
PRE	28.6 (10.1)	16.2 (5.2)	11.5 (4.6)	17.7 (5.5)
POST	22.4 (7.9)	13.5 (4.8)	10.7 (5.7)	19.3 (6.7)
<i>Delay Group</i>				
PRE	25.3 (10.0)	14.3 (4.9)	11.4 (5.1)	19.0 (5.8)
POST	22.9 (11.5)	12.5 (4.9)	11.1 (6.0)	19.6 (5.1)

Table 2: Profile of Mood States results. Values are indicated as mean scores (SD).

### 2.3.d Faces scale

Patients were asked to point at faces to indicate their mood as well as their evaluation of the therapy sessions. We found that patients in the delay group tended to give slightly more positive self-mood ratings during the PRE measurement [Mann-Whitney U,  $Z=-1.75, p=.08$ ] but this difference had vanished during the POST measurement [Mann-Whitney U,  $Z=-.51, p=.61$ ]. Patients' ratings of their own mood became more

positive in the course of therapy [Friedman rank sum test  $\chi^2(11)=23.10$ ,  $p=.02$ ]. Ratings of the therapy sessions themselves remained constant [Friedman rank sum test  $\chi^2(9)=9.15$ ,  $p=.42$ ].

### **2.3.e Auditory and auditory-motor tests**

In the patients of the second phase of our study we measured the purely auditory temporal irregularity detection (anisochrony) both PRE and POST, and the active and passive auditory-motor delay detection task only at POST. For the anisochrony task, we first discarded blocks in which patients responded for more than 30% of the catch (regular) trials that they were irregular. In these cases, the MLP threshold estimate might be unreliable. This was the case for 25% of all recorded blocks. Next, we verified that the blocks had correctly converged at the end by fitting a line to the last 10 trials of each block and discarding blocks in which slope exceeded 1.18 msec/trial. This amounted to discarding an additional 6.7% of blocks. Finally, some blocks yielded a maximal threshold, probably reflecting a misunderstanding of task instructions or incorrect responses. These blocks (6.7%) were again discarded. The number of discarded blocks did not differ between the groups. After this discarding, 4 patients had no valid blocks remaining for the PRE measurement and 3 had no valid blocks remaining for the POST measurement. We restricted our further analysis to patients for whom we had both at least a valid PRE and a valid POST block. This was the case for 8 patients in the delay group and 7 in the normal group. Because of this reduced group size, we preferred nonparametric analyses for paired comparison. For improved readability, we rescaled the threshold as a percentage of the inter-onset-intervals of the tones in the sequence (350 msec). We found no differences between the groups during the PRE measurement [Mann-Whitney U,  $Z=.23$ ,  $p=.81$ ] with an average threshold of 13.8 (SD=8.8) %IOI. There was no difference between the PRE and POST measurements for the normal group [Wilcoxon  $V=16$ ,  $p=.81$ ] or the delay group [Wilcoxon  $V=30$ ,  $p=.11$ ].

For the active delay detection task, we discarded 13% of blocks because of 'delayed' responses to catch trials (which were immediate). Furthermore, 25% of blocks yielded maximal readings and 10% failed to converge. We had to discard 4 patients who had no one valid block remaining and were left with 8 delay and 7 normal group participants. Crucially,

the thresholds for the delay group were higher ( $M=645$ ,  $SD=143$  msec) than for the normal group ( $M=439$ ,  $SD=233$  msec) [Mann-Whitney U,  $Z=2.08$ ,  $p=.04$ ] (Figure 3).

For the passive delay detection task, we discarded 28% of blocks due to more than 30% catch-trial false responses, a further 13% because of non-convergence and 18% because of a maximal read-out. This meant we had to discard 8 patients for whom no valid blocks remained, leaving 8 patients in the delay group and 4 in the normal group. For this reason, we could not pursue data analysis on this test.

The overall finding, then, is that the patients in both groups had equivalent general auditory processing capacity, but during the POST measurement the auditory-motor delay detection threshold was higher in the delay group. We verified the anisochrony test for those patients in whom we had delay detection data, but again we found no effect [Mann-Whitney U,  $Z=0.19$ ,  $p=.85$ ], ruling out a potential explanation of differential patient selection (due to block discarding) in the two cases.

A limitation to the results of this section is that we had to discard a major portion of our threshold data. This is a common finding with patient populations and it limits our conclusions. A further limitation is that we could test the delay detection task only during the POST measurement. Otherwise, patients might have been primed to manipulations in feedback in a way that could have distorted the therapy.

### **2.3.f Within-therapy measure: keyboard keystrokes**

For the keystroke log files in each session (which were recorded only for patients in the second phase), we discarded the notes that were played by the therapist (see methods). Patients played more keystrokes as therapy progressed [ $F(9,162)=8.18$ ,  $p_{GG}<.001$ ,  $\eta_G^2=.12$ ] from an average 247 ( $SD=165$ ) keystrokes in the first therapy session to 459 ( $SD=176$ ) keystrokes in the tenth session. There was a trend for the normal group to play more notes ( $M=489.8$ ,  $SD=194$  keystrokes) than the delay group ( $M=350$ ,  $SD=119$  keystrokes) [ $F(1,18)=4.05$ ,  $p=.06$ ,  $\eta_G^2=.14$ ]. There was no interaction between group and session [ $F(9,162)=.45$ ,  $p_{GG}=.81$ ]. The strength of the keystrokes (measured as MIDI keystroke velocity) remained the same throughout the therapy [ $F(9,162)=.73$ ,  $p_{GG}=.59$ ] and there were no effects of group [ $F(1,18)=.28$ ,  $p=.60$ ] or interaction [ $F(9,162)=.35$ ,  $p_{GG}=.86$ ].

We investigated the scale playing movement (i.e. pressing the five adjacent keys in ascending and then descending order), since this was practiced in every session. We extracted the correctly played scales from each session log-file and then discarded the inter-keystroke-intervals that were further than 3 SD away from the mean for that session. However, not every patient managed to play one iteration of the figure correctly in every session. Therefore we binned the sessions and present the results together for sessions 1-3, 4-6 and 7-10. In this way, we had data in every cell and were able to perform an ANOVA. We performed an ANOVA with within-participant factor session (three levels: sessions 1-3, 4-6 and 7-10) and between-participants factor group (normal, delay). Patients played an approximately 12.4 (SD=6.75) correct scales in one session, and this number did not change over the course of therapy [ $F(2,36)=2.13$ ,  $p=.14$ ] or by group [ $F(1,18)=2.06$ ,  $p=.17$ ] and there was no interaction [ $F(2,36)=1.91$ ,  $p=.16$ ]. We quantified the median inter-keystroke-interval (IKIs) during scale playing (i.e., a measure for the speed of the scale playing) for each session bin and we found that the delay group played significantly slower ( $M=1461$ ,  $SD=502$  msec IKI) than the normal group ( $M=883$ ,  $SD=410$  msec IKI) [ $F(1,18)=7.32$ ,  $p=.01$ ,  $\eta^2=.27$ ]. There was a small effect showing that scale playing became faster in the course of therapy from 1321 (SD=622) msec IKI during sessions 1-3 to 1177 (SD=508) msec IKI in sessions 7-10 [ $F(2,36)=3.71$ ,  $p_{GG}=.045$ ,  $\eta^2=.02$ ] but no interaction with group [ $F(2,36)=.77$ ,  $p_{GG}=.44$ ] (Figure 4). Variability in timing was calculated as the coefficient of variation of the inter-keystroke-intervals and it was on average 33.5 (SD=8.02) % CV-IKI. This quantity did not differ between the groups [ $F(1,18)=0.20$ ,  $p=.66$ ], sessions [ $F(2,36)=.83$ ,  $p=.44$ ] and there was no interaction [ $F(2,36)=0.82$ ,  $p=.45$ ].

### **2.3.g Qualitative interviews**

Qualitative interviews were performed with the patients included in the second phase of this study. The interviews confirmed that patients overall enjoyed the musical aspect of the therapy, and some even indicated that they would continue to practise at home. When asked whether the therapy had been beneficial, reports were mixed. A majority of patients reported that therapy had helped to improve their movements, but a few had the impression there was no change.

Crucially, we tried to establish whether patients in the delay group had noticed the

delay that the keyboard generated. In order to overcome the inhibition patients might have reporting an aspect of the therapy that could be construed as criticism, we tried to use indirect questions. For example, we told the patients we were thinking of buying a different keyboard that would react better to the keystroke and then asked them whether they considered this a good idea. The vast majority of patients responded negatively and said they saw no need to improve the set-up, indicating that they had likely not noticed any delay. An exception was one patient who had clearly noticed the delay: “The keystroke is [...] too slow, it takes too long when you press it until the sound then comes from the speaker.” This patient had a delay detection threshold of 455 msec which was the one-but-lowest in the delay group (the patient with the lowest threshold did not mention noticing a change).

As suggestions for future improvements in the therapy, patients suggested slightly longer sessions, a better visual notation system (including numbered stickers on the keys) or wider keys.

## 2.4 Discussion

The present study tested the hypothesis whether sensory systems can be recruited in motor rehabilitation after stroke, taking music-supported therapy as a test case. In the introduction, we listed three possible explanations for the therapeutic benefit of music-supported stroke rehabilitation. First, patients might benefit from the mere repetition of movements involved in music playing, with the sound having no influence on motor rehabilitation (*movement-only hypothesis*). Alternatively, patients may be motivated by music and therefore perform better (motivation hypothesis). Finally, patients might benefit from feedback about their movement that is provided in the form of sound following each keystroke (auditory-motor hypothesis). Our principal aim was to test the latter hypothesis. The coupling between auditory and motor events was experimentally disrupted in the *delay* group by introducing a random delay between the piano keystroke and the resulting sound. In the normal group, the piano responded immediately. In this way, our aim was provide a

tightly controlled test for involvement of the auditory modality in motor rehabilitation.

Notice that we are interested in (relatively far) transfer. The goal of our therapy is not to teach people music or to make them excellent musicians. Instead, the activity of music playing is taken as a skillful and enjoyable activity whose training is hoped to then transfer to activities of daily activity (such as measured by the nine-hole pegboard test, or, more indirectly, by finger tapping).

Our study differs from previous studies in music-supported therapy in that we restricted the therapy to piano training, so that we could provide a more controlled therapeutic setting. Our study also added a number of measures (anisochrony detection, delay detection, synchronisation-continuation tapping) and methods of analysis (finger tapping regularity) that were not reported in previous clinical trials of music-supported therapy. Furthermore, we monitored finger tapping speed and regularity as well as a simple mood measurement (faces scale) during each individual therapy session, allowing us to pinpoint when during the course of therapy critical improvements happened.

We will now discuss first those tests on which both groups showed equal improvements, and then the tests that revealed differences between the groups. The two groups showed improvement in finger tapping speed. The majority of this improvement happened in the first few sessions of therapy, after which improvements still happen but at a much slower pace. Our study is the first to investigate this improvement in finger tapping over the course of therapy instead of only before and after therapy. This provided crucial information for clinical practice of music-supported therapy as it suggests that when tapping speed is the main improvement goal of therapy (as was the case in previous studies (Amengual et al., 2013; Rojo et al., 2011; Schneider et al., 2007)), then the therapeutic intervention could be shortened without loss in effectiveness. However, it might also be considered that the second part of therapy is important for consolidation. Tapping speed was furthermore shown to slightly improve after the therapy session relative to before. This rules out the possibility that the session had fatigued patients, and, together with patient's subjective reports, leads us to suggest that sessions could be made slightly longer (for example, 45 minutes, although of course the results of this would have to be tested first). Furthermore, patients tapped with less force over the course of therapy, and the dwell (contact) duration during the tapping cycle became longer. That is, the proportion of the tap

cycle during which the finger is in contact with the table is longer. No differences were found between the groups. The fact that patients tapping becomes less forceful might be a sign of reduced muscle co-activation, less spasticity and therefore improved motor control (Alon, Sunnerhagen, Geurts, & Ohry, 2003).

Furthermore, patients in both groups improved in mood, as evidenced by a marked reduction in depression/anxiety and fatigue. This finding is encouraging in the light of the relative high incidence of depression among acute stroke patients (Kouwenhoven, Gay, Bakken, & Lerdal, 2013). However, since these mood improvements do not differ between the groups, they cannot strictly be attributed to our therapeutic intervention (although our result shows that our sound manipulation did not affect these improvements). Instead, they could also be effects of other therapies the patients receive in parallel to their participation in our study, or they could be a general result of being in a rehabilitation setting. For this reason, our emphasis is on measures that revealed a differential effect in the two groups (normal and delay).

Now we turn to the key differences between the normal and delay groups. We hypothesised that fine motor rehabilitation would be less effective in the group receiving a distorted (jittered delay) auditory feedback. Contrary to our expectations, patients in the delay group improve *more* than those in the *normal* group, as evidenced, for example, by the Nine-hole pegboard test. Further, patients in the delay group exhibited greater regularity in their finger tapping than did the patients in the normal group. Surprisingly, the difference between the groups disappeared during the POST measurement because of a sudden improvement in tapping regularity in the normal group. The delay group, however, maintains the same level of regularity at this point. First of all, note that the difference in tapping regularity cannot be explained by a direct influence of the auditory feedback. Patients were not using the piano but were tapping on the table (and force sensor) and therefore did not receive any altered feedback (nor, actually, any feedback at all apart from the sound their finger made on the table, the tactile sensation of impact and proprioception).

We will now discuss in detail possible explanations for the finding that the delay group outperforms the normal group in terms of fine motor control (Nine-hole pegboard test) and tapping regularity. Firstly, we point out that the finding provides evidence against

the interpretation that it is only the movement (and not the sound) that contributes to motor rehabilitation (the *movement-only explanation*). The reason is that the only difference between our groups was the way in which the keyboard that was used during therapy emitted the sounds. Our findings support the idea that crosstalk between the auditory and motor systems plays a role in the rehabilitation (the *auditory-motor explanation*).

Our prediction, based on the auditory-motor explanation, was that the normal group would outperform the delay group. If patients' motor re-learning is based on a sensory feedback model in which the feedback was auditory (instead of proprioceptive or visual), then increasing the (temporal) noise in the feedback signal should disrupt learning. The finding was the opposite. Several explanations come to mind, which we will discuss here, although we point out that since these interpretations are all post-hoc, they should be treated with caution until targeted studies provide more direct tests.

A first possible explanation is that patients use an auditory feedback model, but those in the delay group benefit from the fact that its predictions of sensory consequences would mostly be wrong. The reasoning is as follows. The brain generates temporal predictions of sensory consequences of self-generated movements, possibly using forward internal models (Wolpert & Kawato, 1998). In the delay feedback condition, such predictions will continue to mostly be wrong, since the delay in the sound is randomly sampled at every keystroke and therefore cannot be accurately predicted. As a result, the brain will update its internal models continually, contrary to patients in the normal condition who will leave their internal model mostly as is. We speculate that the updating of the internal model may induce a certain flexibility in the motor representations that would then be beneficial for motor control itself (Assaiante, Chabeauti, Sveistrup, & Vaugoyeau, 2011). The interpretation is supported by our finding that the delay group's auditory-motor delay detection threshold is higher than that of the normal group, suggesting that they have to an extent adapted to the delay by shifting their subjective auditory-motor synchrony (Sugano et al., 2010). In order to test this interpretation, a future study could contrast a group of patients playing on a constantly delayed piano with those playing on a jittered delay piano (as in our study). If the delay is constant, the brain is able, at least in principle, to generate accurate predictions of when the sensory effect (tone) will happen relative to each keystroke. This interpretation would therefore predict that the

jittered delay group would outperform the constant delay group, because the latter group does not benefit from continual updating of their sensory predictions.

A second explanation is that patients in the delay group receive a greater error signal based on their playing. Simply put, when they play as regularly as the patients in the normal group, the sounds come out less regular nevertheless. Therefore, the delay group may have a greater motivation to improve their playing. In other words, they would work harder. Although we cannot test this directly, it may be less likely that this is the case. This account would predict that patients in the delay group would enjoy therapy less (because they are unhappy with the auditory result of their movements). However, we found no mood or therapy session evaluation differences between the groups. Also, we would have expected that the delay group would be motivated to play more and take less breaks, but instead we found there was a tendency for the normal group to play more during the therapy sessions.

A third explanation suggests that in stroke patients the auditory-motor system is damaged and therefore auditory sensorimotor integration is compromised (Rodriguez-Fornells et al., 2012). As a result, the auditory feedback that benefits healthy populations actually becomes a distracter by providing incorrect or noisy information or making the patient dependent on it (Salmoni, Schmidt, & Walter, 1984). Under this interpretation, auditory feedback provides a lure by attracting patient's attention whereas its actual effects on movement are counterproductive. This phenomenon is similar to stutterers or patients with conduction aphasia. These populations show a detrimental effect of immediate auditory feedback, probably because the mechanism by which they integrate this feedback in their movements is noisy (Civier, Tasko, & Guenther, 2010; Hickok, Houde, & Rong, 2011; Max, Guenther, Gracco, Ghosh, & Wallace, 2004). In the delay feedback group, the patients' brain implicitly notices the high uncertainty associated with the auditory-motor mapping and is pushed to discard it in favour of relatively more secure sensory inputs such as vision, proprioception or tactile information. Indeed, it is a common finding that the brain weights the input from multiple sensory channels in a Bayesian optimal way (Kording & Wolpert, 2004), which in the case of jittered delayed feedback would predict reduction of reliance on auditory information. This account can explain the sudden improvement of the normal group in tapping regularity between the tenth session

and the POST measurement, for the short-term counterproductive effects of the immediate-feedback music training have worn off. If this explanation is correct, it would mean we need to re-think our interpretation of music-supported therapy, since it would imply that in some cases the auditory information can be a burden instead of a help. Instead, jittered delayed feedback might provide a skilful way to allow patients to benefit from positive motivational effects of music (Thompson et al., 2001) without the potentially counterproductive effects of reliance on feedback. This interpretation is in line with previous interpretations that suggested malfunctioning auditory-motor connections (Rodriguez-Fornells et al., 2012). A prediction derived from our hypothesis that can be directly tested is that we expect less or no auditory-motor co-activation (Rojo et al., 2011) in pianists in the delay feedback group.

Whichever of the explanations advanced here is correct (or yet another future explanation) can be the topic of future research. Independent of this question, there appear to be a number of conclusions that are common to these explanations. Firstly, it seems patients have continued to associate the sounds with their movements, likely through a larger temporal integration window (which is a common construct in multisensory integration, see e.g. Stevenson & Wallace, 2013) or through shifting of auditory-motor synchrony (Sugano et al., 2010). It is difficult to account for these data by assuming that auditory information plays no role whatsoever in music-supported therapy.

Finally, we turn to the interesting finding that patients responded to the delayed auditory feedback by playing slower. The finding squares with research in musicians who also slow down with delayed auditory feedback (Pfordresher & Dalla Bella, 2011; Pfordresher & Palmer, 2002). Therefore, this finding points us to an aspect of auditory-motor functioning in patients that is not impaired, but normal. Furthermore, this finding raises the additional explanation that performing movements more slowly might improve motor rehabilitation learning. Future studies could investigate this question by having different groups of patients play music at different speeds and investigating whether they show a differential motor outcome.

A limitation of the present study is that we used liberal inclusion criteria instead of selecting patients by some precise lesion sites in the brain. Another limitation of our study is that we did not employ a control condition that did not receive music rehabilitation at all.

Therefore, changes that do not differ between our experimental groups (such as the reduction in depression/anxiety and fatigue) cannot strictly be conclusively attributed to our musical intervention. They may be a general effect of the various other therapies that patients participated in parallel to our study. However, the advantage of our approach is that it provides the best controlled test for involvement of auditory-motor coupling in stroke rehabilitation. Furthermore, a previous study has compared music-supported therapy with other therapies, although not in a randomised fashion (Schneider et al., 2010), making it likely that the improvements that we see across groups can be attributed to music-supported therapy. We propose that future studies may use similar manipulations of single aspects of music therapy as we have done in order to better understand the active ingredients. In this way, music-supported therapy can serve as its own control condition. Understanding what mechanisms cause motor improvements in music-supported therapy in turn will allow optimisation of clinical implementations with maximal benefit to the patients.

## 2.5 References

- Alon, G., Sunnerhagen, K. S., Geurts, A. C. H., & Ohry, A. (2003). A home-based, self-administered stimulation program to improve selected hand functions of chronic stroke. *NeuroRehabilitation*, 18(3), 215–225.
- Altenmüller, E., Marco-Pallares, J., Münte, T. F., & Schneider, S. (2009). Neural reorganization underlies improvement in stroke-induced motor dysfunction by music-supported therapy. *Annals of the New York Academy of Sciences*, 1169, 395–405. doi:10.1111/j.1749-6632.2009.04580.x
- Amengual, J. L., Rojo, N., Veciana de las Heras, M., Marco-Pallarés, J., Grau-Sánchez, J., Schneider, S., ... Rodríguez-Fornells, A. (2013). Sensorimotor Plasticity after Music-Supported Therapy in Chronic Stroke Patients Revealed by Transcranial Magnetic Stimulation. *PLoS ONE*, 8(4), e61883. doi:10.1371/journal.pone.0061883
- Andrews, F. M., & Crandall, R. (1976). The validity of measures of self-reported well-being. *Social Indicators Research*, 3(1), 1–19. doi:10.1007/BF00286161
- Assaiante, C., Chabeauti, P.-Y., Sveistrup, H., & Vaugoyeau, M. (2011). Updating process of internal model of action as assessed from motor and postural strategies in young adults. *Human movement science*, 30(2), 227–237. doi:10.1016/j.humov.2010.05.007
- Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision Substitution by Tactile Image Projection. *Nature*, 221(5184), 963–964. doi:10.1038/221963a0
- Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neuroscience*, 4, 26–26. doi:10.1186/1471-2202-4-26
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., ... Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, 30(3), 917–926. doi:10.1016/j.neuroimage.2005.10.044
- Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research*, 1161, 65–78. doi:10.1016/j.brainres.2007.05.045
- Bernier, P.-M., Chua, R., Bard, C., & Franks, I. M. (2006). Updating of an internal model without proprioception: a deafferentation study. *Neuroreport*, 17(13), 1421–1425. doi:10.1097/01.wnr.0000233096.13032.34
- Bullinger, M., Heinisch, M., Ludwig, M., & Geier, S. (1990). Skalen zur Erfassung des Wohlbefindens: Psychosomatische Analysen zum “Profile of Mood States” (POMS) und zum “Psychological General Wellbeing Index” (PGWI). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 11(1), 53–61.
- Calautti, C., Jones, P. S., Persaud, N., Guincestre, J.-Y., Naccarato, M., Warburton, E. A., & Baron, J.-C. (2006). Quantification of index tapping regularity after stroke with tri-axial accelerometry. *Brain Research Bulletin*, 70(1), 1–7. doi:10.1016/j.brainresbull.2005.11.001
- Carey, L. M., Oke, L. E., & Matyas, T. A. (1996). Impaired limb position sense after stroke: A quantitative test for clinical use. *Archives of Physical Medicine and Rehabilitation*, 77(12), 1271–1278.
- Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of fluency disorders*, 35(3), 246–279. doi:10.1016/j.jfludis.2010.05.002
- Connell, L., Lincoln, N., & Radford, K. (2008). Somatosensory impairment after stroke: frequency of different deficits and their recovery. *Clinical Rehabilitation*, 22(8), 758–767. doi:10.1177/0269215508090674
- Curran, S. L., Andrykowski, M. A., & Studts, J. L. (1995). Short Form of the Profile of Mood States (POMS-SF): Psychometric information. *Psychological Assessment*, 7(1), 80.
- Dimyan, M. A., & Cohen, L. G. (2011). Neuroplasticity in the context of motor rehabilitation after stroke. *Nat*

*Rev Neurol*, 7(2), 76–85. doi:10.1038/nrneurol.2010.200

- Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: influence of the tempo and musical backgrounds of listeners. *Brain and Cognition*, 58(1), 133–147. doi:10.1016/j.bandc.2004.09.014
- Fisher, N. I. (1995). *Statistical Analysis of Circular Data*. Cambridge University Press.
- Grau-Sánchez, J., Amengual, J. L., Rojo, N., Heras, M. V. de las, Rubio, F., Altenmüller, E., ... Rodríguez-Fornells, A. (2013). Plasticity in the sensorimotor cortex induced by Music-supported therapy in stroke patients: a TMS study. *Frontiers in Human Neuroscience*, 7, 494. doi:10.3389/fnhum.2013.00494
- Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes–no task. *The Journal of the Acoustical Society of America*, 93(4), 2096. doi:10.1121/1.406696
- Gu, X., & Green, D. M. (1994). Further studies of a maximum-likelihood yes–no procedure. *The Journal of the Acoustical Society of America*, 96(1), 93. doi:10.1121/1.410378
- Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of cognitive neuroscience*, 17(2), 282–293. doi:10.1162/0898929053124893
- Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, 13(6), 786–792. doi:10.1162/08989290152541449
- Heitger, M. H., Ronsse, R., Dhollander, T., Dupont, P., Caeyenberghs, K., & Swinnen, S. P. (2012). Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis. *NeuroImage*, 61(3), 633–650. doi:10.1016/j.neuroimage.2012.03.067
- Heller, A., Wade, D. T., Wood, V. A., Sunderland, A., Hower, R. L., & Ward, E. (1987). Arm function after stroke: measurement and recovery over the first three months. *Journal of Neurology, Neurosurgery, and Psychiatry*, 50(6), 714–719.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron*, 69(3), 407–422. doi:10.1016/j.neuron.2011.01.019
- Hove, M. J., & Keller, P. E. (2010). Spatiotemporal Relations and Movement Trajectories in Visuomotor Synchronization. *Music Perception: An Interdisciplinary Journal*, 28(1), 15–26.
- Hyde, K. L., & Peretz, I. (2004). Brains That Are Out of Tune but in Time. *Psychological Science*, 15(5), 356–360. doi:10.1111/j.0956-7976.2004.00683.x
- Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244–247. doi:10.1038/nature02169
- Kouwenhoven, S. E., Gay, C. L., Bakken, L. N., & Lerdal, A. (2013). Depressive symptoms in acute stroke: A cross-sectional study of their association with sociodemographics and clinical factors. *Neuropsychological rehabilitation*. doi:10.1080/09602011.2013.801778
- Kunin, T. (1955). The Construction of a New Type of Attitude Measure. *Personnel Psychology*, 8, 65–77. doi:10.1111/j.1744-6570.1955.tb01189.x
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired Actions. *The Journal of Neuroscience*, 27(2), 308–314. doi:10.1523/JNEUROSCI.4822-06.2007
- Lorr, M., McNair, D., & Droppleman, L. (1971). Manual: profile of mood states. *San Diego, CA: Educational and Industrial Testing Service*.
- Mathiowetz, V., Kashman, N., Volland, G., Weber, K., Dowe, M., & Rogers, S. (1985). Grip and pinch strength: normative data for adults. *Archives of Physical Medicine and Rehabilitation*, 66(2), 69–74.
- Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary issues in communication science and disorders*, 31, 105–122.
- McDowell, I. (2006). *Measuring health: a guide to rating scales and questionnaires*. Oxford University Press.
- Okuno, R., Yokoe, M., Akazawa, K., Abe, K., & Sakoda, S. (2006). Finger taps movement acceleration measurement system for quantitative diagnosis of Parkinson's disease. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE*

*Engineering in Medicine and Biology Society. Conference, Suppl*, 6623–6626.  
doi:10.1109/IEMBS.2006.260904

- Oscari, F., Secoli, R., Avanzini, F., Rosati, G., & Reinkensmeyer, D. J. (2012). Substituting auditory for visual feedback to adapt to altered dynamic and kinematic environments during reaching. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 221(1), 33–41.  
doi:10.1007/s00221-012-3144-2
- Parker, V. M., Wade, D. T., & Langton Hewer, R. (1986). Loss of arm function after stroke: measurement, frequency, and recovery. *International Rehabilitation Medicine*, 8(2), 69–73.
- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and Emotions in the Brain: Familiarity Matters. *PLoS ONE*, 6(11), e27241. doi:10.1371/journal.pone.0027241
- Pfordresher, P., & Dalla Bella, S. (2011). Delayed auditory feedback and movement. *Journal of Experimental Psychology. Human Perception and Performance*, 37(2), 566–579. doi:10.1037/a0021487
- Pfordresher, P., & Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research*, 66(1), 71–79.
- Rodriguez-Fornells, A., Rojo, N., Amengual, J. L., Ripollés, P., Altenmüller, E., & Münte, T. F. (2012). The involvement of audio-motor coupling in the music-supported therapy applied to stroke patients. *Annals of the New York Academy of Sciences*, 1252(1), 282–293. doi:10.1111/j.1749-6632.2011.06425.x
- Rojo, N., Amengual, J., Juncadella, M., Rubio, F., Camara, E., Marco-Pallares, J., ... Rodriguez-Fornells, A. (2011). Music-Supported Therapy induces plasticity in the sensorimotor cortex in chronic stroke: A single-case study using multimodal imaging (fMRI-TMS). *Brain Injury*.  
doi:10.3109/02699052.2011.576305
- Ronsse, R., Puttemans, V., Coxon, J. P., Goble, D. J., Wagemans, J., Wenderoth, N., & Swinnen, S. P. (2011). Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. *Cerebral cortex (New York, N.Y.: 1991)*, 21(6), 1283–1294. doi:10.1093/cercor/bhq209
- Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychological bulletin*, 95(3), 355–386.
- Särkämö, T., Tervaniemi, M., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., ... Hietanen, M. (2008). Music listening enhances cognitive recovery and mood after middle cerebral artery stroke. *Brain*, 131(3), 866–876. doi:10.1093/brain/awn013
- Schabrun, S. M., & Hillier, S. (2009). Evidence for the retraining of sensation after stroke: a systematic review. *Clinical Rehabilitation*, 23(1), 27–39. doi:10.1177/0269215508098897
- Schneider, S., Münte, T., Rodriguez-Fornells, A., Sailer, M., & Altenmüller, E. (2010). Music-Supported Training is More Efficient than Functional Motor Training for Recovery of Fine Motor Skills in Stroke Patients. *Music Perception*, 27(4), 271–280.
- Schneider, S., Schönle, P. W., Altenmüller, E., & Münte, T. F. (2007). Using musical instruments to improve motor skill recovery following a stroke. *Journal of Neurology*, 254(10), 1339–1346.  
doi:10.1007/s00415-006-0523-2
- Sherrington, C. S. (1907). On the proprioceptive system, especially in its reflex aspect. *Brain*, 29(4), 467–482.  
doi:10.1093/brain/29.4.467
- Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 227(2), 249–261.  
doi:10.1007/s00221-013-3507-3
- Sugano, Y., Keetels, M., & Vroomen, J. (2010). Adaptation to motor-visual and motor-auditory temporal lags transfer across modalities. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 201(3), 393–399. doi:10.1007/s00221-009-2047-3
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, 12(3), 248–251.
- Vidoni, E. D., & Boyd, L. A. (2009). Preserved motor learning after stroke is related to the degree of proprioceptive deficit. *Behavioral and Brain Functions: BBF*, 5, 36. doi:10.1186/1744-9081-5-36
- Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses.

*Perception & Psychophysics*, 14(1), 5–12. doi:10.3758/BF03198607

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks: The Official Journal of the International Neural Network Society*, 11(7-8), 1317–1329.

Wong, J. D., Kistemaker, D. A., Chin, A., & Gribble, P. L. (2012). Can proprioceptive training improve motor learning? *Journal of neurophysiology*, 108(12), 3313–3321. doi:10.1152/jn.00122.2012

## 2.6 Conflict of interest statement

The authors declare this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## 2.7 Acknowledgements

This work was supported by the EBRAMUS, European Brain and Music Ph.D. Grant to FTV (ITN MC FP7, GA 238157). We are indebted to Britta Westner, M.A., for running a pilot study. We furthermore would like to thank Mr. Richter who helped us test the therapy. Also, Dr. Sabine Schneider kindly shared her knowledge of the previous implementation of the music-supported therapy that formed the basis for the current therapy program. We thank all the medical staff (doctors and nurses) in the Hessisch Oldendorf clinic for their cooperation and for indicating to us patients who might be suitable for inclusion in the study. Finally, but not least, we wish to extend our gratitude to all the patients who devoted their time to participation in this study. We hope the therapy may benefit them in their daily lives.

Sounds on time - Re-learning in stroke patients

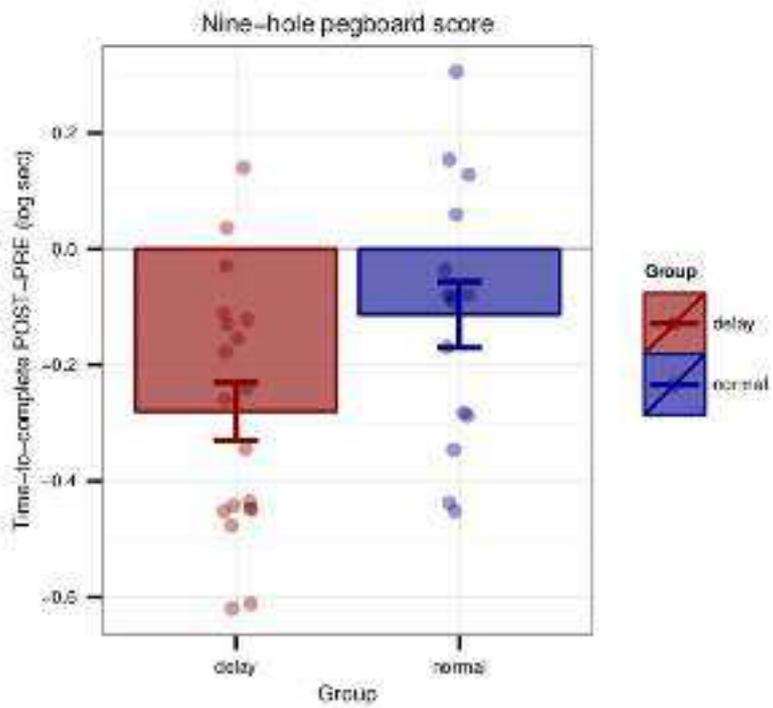


Figure 1: Nine hole pegboard result. Error bars indicate the standard error of the mean and the dots represent the individual patients.

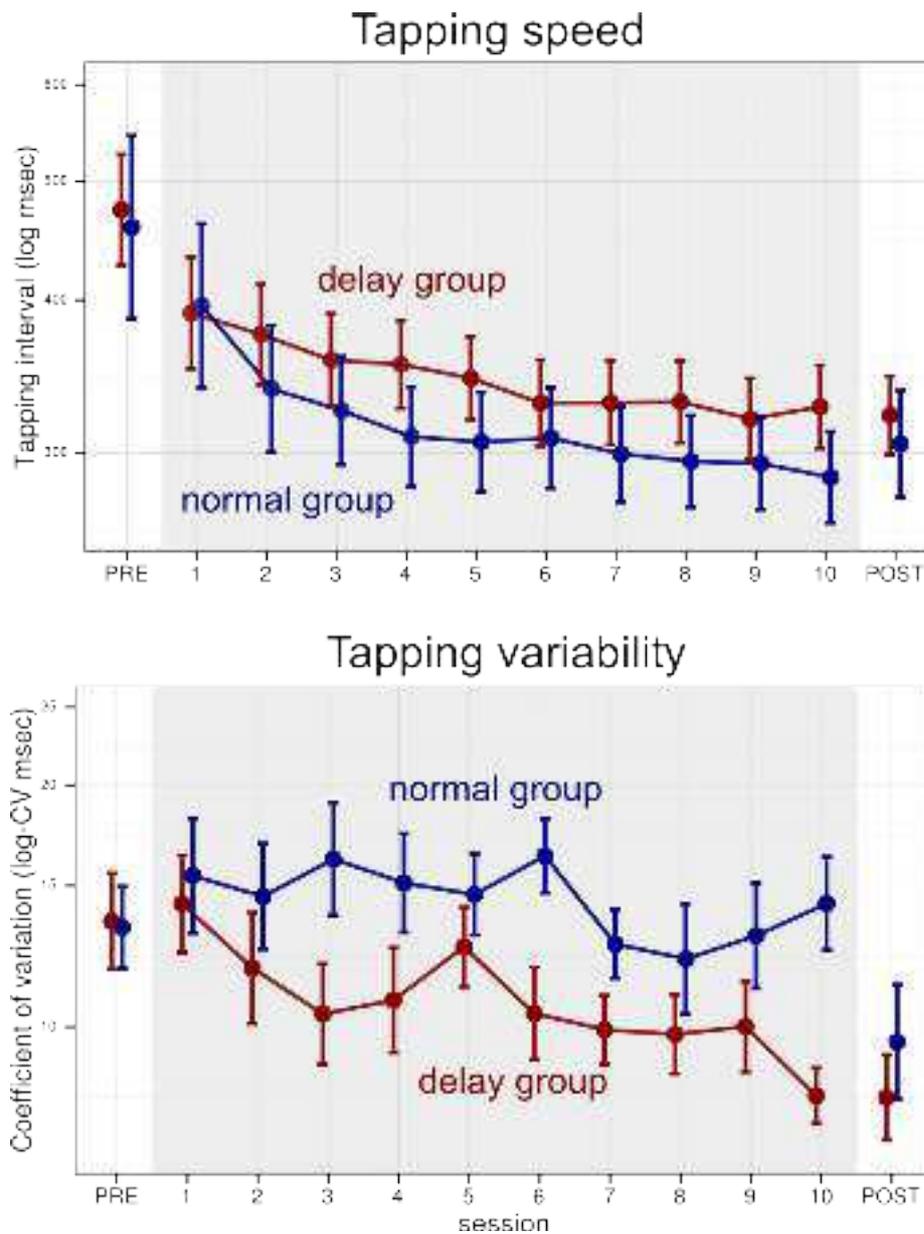


Figure 2: Tapping speed (top) and variability (bottom) of the index finger before, during and after therapy. The x-axis indicate the PRE measurement, the ten therapy sessions and the POST measurement. Dots indicate the mean for the two groups and error bars indicate standard error of the mean. Variability is expressed as coefficient-of-variation (CV). A higher CV value indicates more variable tapping, a lower CV indicates less variable, more regular tapping.

Sounds on time - Re-learning in stroke patients

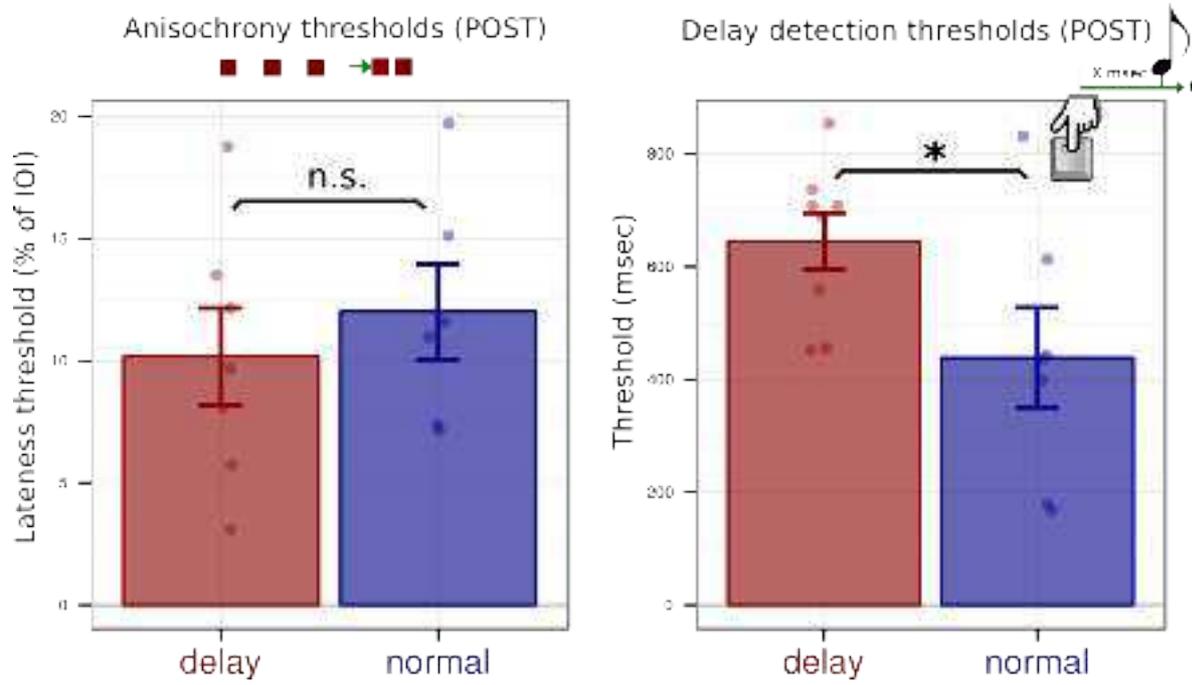


Figure 3: Auditory (anisochrony) and auditory-motor (delay detection) threshold results.

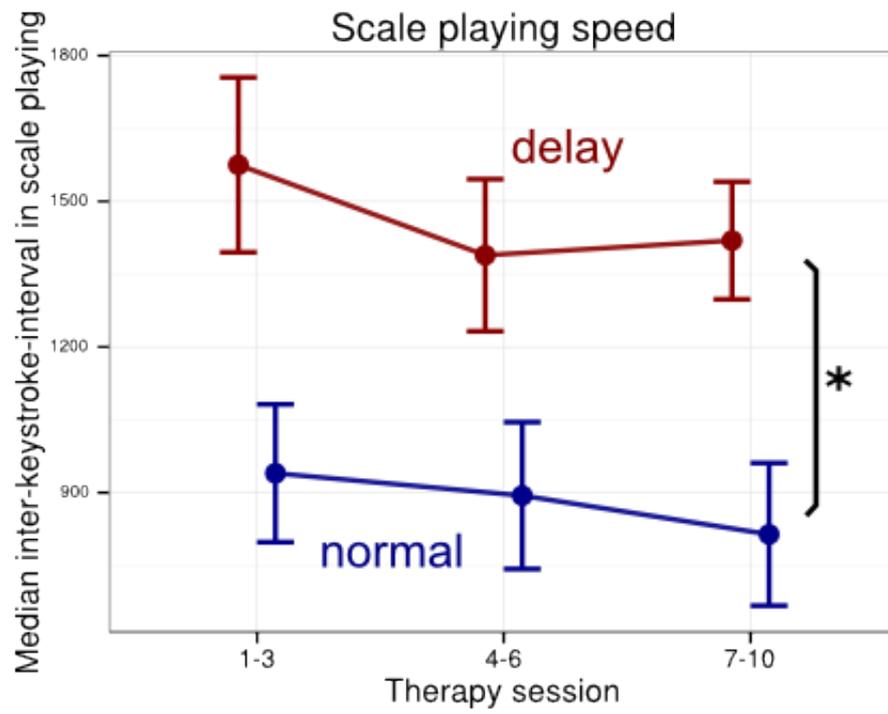


Figure 4: Scale playing speed for the two groups and across therapy.

### **III.3. Music-supported motor training after stroke reveals no superiority of synchronisation in group therapy**

---

## Music-supported motor training after stroke reveals no superiority of synchronisation in group therapy

van Vugt FT<sup>1,2\*</sup>, Ritter J<sup>1\*</sup>, Rollnik J<sup>3</sup>, Altenmüller E<sup>1</sup>

\* shared first-authorship

<sup>1</sup>Institute of Music Physiology and Musicians' Medicine, University of Music, Drama and Media, Hannover, Germany

<sup>2</sup>Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, CNRS-UMR 5292, INSERM U1028, University Lyon-1, Lyon, France

<sup>3</sup>BDH-Klinik, Hessisch Oldendorf, Germany

### Abstract

*Background.* Motor impairments are among the most common and most disabling consequences of stroke. Recently, patients showed significant improvement in fine motor skills after music-supported therapy in which patients learned to play the piano with the paretic hand. A unique feature of music is that it is inherently social: music can be played together as no other activity can be done collectively. In particular, musical movements can be performed in synchrony.

*Aim.* We explored the potential of synchronised music playing during therapy and whether it could benefit fine motor rehabilitation and mood.

*Method.* Twenty-eight patients in early stroke rehabilitation with no previous musical background learned to play simple finger exercises and familiar children's songs on the piano for ten sessions of half an hour. Patients first received three individual therapy sessions and then continued in pairs. The patients were divided into two groups. Patients in one group played together (together group) whereas the patients in the other group played one after the other (in-turn group). To assess fine motor skill recovery the patients performed standard clinical tests such as the nine-hole-pegboard test (9HPT) and index finger-tapping speed and regularity, and metronome-paced finger tapping. Patients' mood was established using the Profile of Mood States (POMS).

*Results.* Both groups showed better fine motor control in the nine-hole-pegboard test, subtle

improvements in index finger tapping speed. Furthermore, in metronome-paced finger tapping, patients in both groups improved significantly. Mood tests revealed equal reductions in both groups in depression and fatigue, whereas no change was found in vigor and hostility. During therapy, we found that patients in the in-turn group rated their partner as more sympathetic than the together-group in a visual-analogue scale.

*Conclusions.* Our results reveal that music-supported stroke rehabilitation can be effective not only individually but also in patient pairs. Patients improved in fine motor skills as well as mood. Patients who were playing in turn rather than simultaneously revealed greater improvements in fine motor skill. We speculate that patients in the former group may benefit from the opportunity to learn from observation.

### 3.1 Introduction

Motor impairments are among the most common and most disabling consequences of stroke (Dimyan & Cohen, 2011; Ward & Cohen, 2004). Recently, patients showed significant improvements in fine motor control after music-supported therapy in which they learned to play the piano and drums during several weeks (Altenmüller, Marco-Pallares, Münte, & Schneider, 2009; Amengual et al., 2013; Grau-Sánchez et al., 2013; Schneider, Schönle, Altenmüller, & Münte, 2007). The researchers found that their musical training transferred to benefits in a range of clinical tasks of activities of daily living, revealing gains in fine motor control in these patients.

A variety of explanations has been advanced for the performance improvement due to music-supported therapy in these patients, such as auditory feedback, novelty of the intervention, and increased motivation (Altenmüller et al., 2009). However, we wondered whether the fact that music is a social activity played a role in music-supported therapy. In healthy populations, music turned out to be an effective tool for supporting pro-social commitment, increasing group cohesion and cooperation (Overy, 2012). The particular aspect of music that appears to be involved in creating group cohesion is synchronisation. Performing (musical) movements in synchrony has been shown to improve feelings of reciprocal likeability (Hove & Risen, 2009), trust (Launay, Dean, & Bailes, 2013; Wiltermuth

& Heath, 2009), pseudo-altruism (Kokal, Engel, Kirschner, & Keysers, 2011; Valdesolo & Desteno, 2011) and even destructive obedience (Wiltermuth, 2012). As a result of these findings, we hypothesised that music may be a powerful tool for creating pro-social engagement in a rehabilitation therapy session. In support of this, it has been shown that participants in music therapy rehabilitation are more actively involved and cooperative than in other forms of therapy (Narme et al., 2012).

Furthermore, a greater engagement in rehabilitation is associated with an improved clinical outcome. Stroke victims often suffer from disturbances in motivation and mood (Caeiro, Ferro, & Costa, 2013). Social support (Sandin, Cifu, & Noll, 1994), in turn, has been associated with improved functional outcome after stroke (Glass, Matchar, Belyea, & Feussner, 1993). However, to date, no studies connected these two causalities to directly test whether increased social engagement could enhance functional motor outcome after stroke.

The present study aimed to test the potential for music as a tool to create pro-social engagement on the part of patients. In particular, we hypothesised that the aspect of music that might boost social participation on the part of the patient is synchronisation of musical playing. That is, do patients benefit from producing sounds in synchrony? In order to specifically test for the effect of synchronisation whilst keeping other factors constant, we divided our patient population into pairs. Some pairs played in synchrony during their therapy whilst others played in turn.

We asked the following questions. First, we asked whether playing together in synchrony is associated with changes in functional motor outcome. Secondly, we asked whether playing in synchrony or in turn influenced patient's mood. Thirdly, we asked whether patient's basic auditory-motor functioning (such as synchronising to a metronome) was influenced by playing in synchrony or in turn.

### **3.2 Methods**

We assigned patients to one of two groups in a randomised design. Both groups received music therapy in pairs and they played the same selection of finger exercises and children's songs. Patients received ten such therapy sessions. Of these ten therapy sessions,

the first three were individual and the last seven were in pairs. The patients were divided into two groups. Patients in one group played in synchrony (together group) whereas the patients in the other group played one after the other (in-turn group). All patients received therapy in groups of two. Prior to therapy (PRE) and after therapy (POST), all patients completed a battery of tests described below. In between the three individual sessions and the seven joint sessions, we included a small session of measurements (INTER).

### **3.2.a Patient group characteristics**

We aimed at obtaining a representative sample of patients from the hospital population. Consequently, we were not able to maintain high homogeneity of patient selection. However, we feel that by making this choice, our results are maximally relevant to clinical practice. Inclusion criteria were:

- light to moderate motor impairment in the upper extremity due to ischaemia or haemorrhage;
- having residual movement capacity enabling the patient to move the arm and the index finger of the affected hand independently;
- age between 30 and 75 years;
- Barthel Index at least 50;
- right-handedness;
- at most 5 months post-stroke;
- being able to understand and agree with informed consent to participate.

Exclusion criteria were:

- previous musical training for more than 4 years;
- psychiatric problems (as assessed by standard clinical investigation);
- cognitive impairment or aphasia.

Basic clinical data of our patients are listed in Table 1.

	together	in-turn	Statistical comparison
Number of patients	14	14	
Sex (female/male)	6/8	10/4	Fisher exact test p=0.25
Age (years)	65.6 (10.5)	67.1 (11.8)	W=85.0, p=0.56
Stroke area (unknown/medial/posterior/anterior)	4/7/2/1	6/7/1/0	$\chi^2(3) = 1.73$ , p = 0.63
Stroke type (ischemia/haemorrhage)	12/2	12/2	Fisher exact test p=1
Handedness (R/L)	14/0	13/1	Fisher exact test p=1
Affected hand (R/L)	9/5	8/6	Fisher exact test p=1
Barthel index PRE	48.2 (15.0)	48.9 (11.5)	W=96.5, p=0.95
Barthel index POST	72.1 (14.4)	67.7 (14.8)	W=104.0, p=0.54
Days since stroke (at PRE, days)	40.6 (25.6)	45.6 (29.9)	W=89.5, p=0.71
Therapy duration (days)	18.2 (3.0)	18.4 (5.1)	W=96.0, p=0.94
Faces scale mood rating PRE	2.42 (1.22)	1.85 (1.23)	W=123.5, p=0.22

Table 1: Biographical data of the two patient groups. Continuous data are reported as mean (standard deviation).

### 3.2.b Music training

Patients received 10 sessions of half an hour of piano training. The sessions took place on different days over the course of three to four weeks. The day before the first session and a day after the last session was dedicated to the PRE and POST measurements, respectively, which are described in more details below.

The therapy program followed the same structure every day. In the beginning, patients played simple finger exercises such as a five-tone scale up and down and other patterns with their paretic hand. Then patients learned to play one from a set of simple children's songs. If patients reached a sufficient level on one of the songs, they would be invited to learn additional songs from the set prepared by the therapist.

Each patient in the pair played on a M-Key V2 MIDI controller keyboard that was

chosen for its light touch. The two keyboards were connected through the M-Audio Midisport Uno MIDI-to-USB converter to a Linux laptop. The laptop ran a custom made C program that recorded the MIDI events to a file and forwarded it to the software synthesiser Fluidsynth, which generated the sounds using a Steinway sound font. The program additionally changed the MIDI velocity value (loudness) to its maximum value. As a result, all sounds sounded maximally loud, which was done to keep patients' typically very soft keystrokes from being heard. The sounds were then played through Creative Inspire T10 speakers (Creative Labs, Inc.) at a comfortable loudness level. The five keyboard keys used in the therapy were numbered. Songs and exercises were then written in a simplified musical notation as numbers in tabular form and presented visually to the patients as a memory aid.

Patients played the piano with the hand of their affected extremity only. The therapist stood next to the patient and supported the patient's arm when so required. The patient was always encouraged to make as much of the movements by him- or herself as possible. For those patients who were more severely affected, the therapist initially pointed to the fingers or moved them gently, encouraging the patient to make the movements unassisted on the next trial. Throughout therapy, the therapist's role was one of supporting and the aim was always to allow the patient to function as independently as possible instead of dependent on the therapist.

### **3.2.c Nine-hole pegboard test**

The nine-hole pegboard test (9HPT) is a clinical test to assess fine motor control. The patient needs to place 9 small sticks (pegs) in 9 holes and take them out again (Heller et al., 1987; Mathiowetz, Volland, Kashman, & Weber, 1985; Parker, Wade, & Langton Hewer, 1986). The patient is seated comfortably with his or her affected arm resting on the table. The 9HPT board is placed so that he or she can reach it easily, with the side with the peg container at the side of the tested arm. The instruction is to take the pegs one by one from the board and put them in the holes, and then take them out and put them back, again one by one. The experimenter holds a stopwatch that is started once the patient touches the first peg, and is stopped once the patient releases the last peg. This test is performed once at PRE and once at POST.

### 3.2.d Finger tapping measurements

We investigated patient's finger tapping performance of the affected hand as a measure of fine motor control. Three different tapping conditions were measured: (1) paced thumb-to-index tapping, (2) index finger speed tapping, (3) middle finger speed tapping. The tests are described in detail in what follows. In all conditions, patients were seated comfortably at a table on which they rested their arm. In order to have a portable, flexible and yet maximally accurate measurement of finger tapping performance, we custom designed a measurement device. Finger motion was recorded by a triaxial accelerometer (ADXL 335) attached gently to the patient's index or middle finger tip (depending on the task). Tap contact was measured by a force sensitive resistor (FSR SEN09375), which is a small sheet whose electrical resistance changes upon contact in a way that is related to the contact force. Both sensors were read out by an Arduino Duemilanove experimentation board running a custom C program to sample sensors at 3 kHz. The data was then transferred online to a Linux laptop running a custom Python program interfacing through USB and allowing the therapist to preview the data. We made the blueprints of the device set up as well as the custom programs available freely online for future research groups to use (<http://github.com/florisvanvugt/immotion>).

In paced thumb-to-index tapping, patients were instructed to tap as regularly as possible in time with a metronome at 69 BPM (i.e. 1.15 Hz) during 60 seconds from the first tap (Calautti et al., 2006). The metronome sound was generated using direct digital synthesis (DDS) by the Arduino experimental board. Essentially, we created a wave table (440 Hz, 20 msec) which was written to a PWM pin connected to an audio jack plug. A set of Creative Inspire T10 loudspeakers (Creative Labs, Inc.) were plugged into this connector. The patient was instructed to tap as follows. The side of the hand (at the little finger) rested lightly on the table and the fingers were held in a relaxed posture (not at maximum flexion or extension). The index finger and thumb moved to touch each other and then move apart again to a distance of about 5cm (but at least 2cm). The thumb-to-index tapping movement was chosen because it was previously argued to be more natural and a more reliable reflection of activities of daily living (Okuno, Yokoe, Akazawa, Abe, & Sakoda, 2006).

In index finger speed tapping, we measured the maximum tapping rate and

variability during approximately 14 seconds (measured from the first finger tap). The patients were invited to choose where it was most comfortable to tap on the table. His or her elbow rested on the table and the hand lies palm down on the table. The fingers were held in a relaxed posture close to maximal extension but slightly bent so that the position could be sustained without muscular effort. No metronome was used in these speed tapping. The patient was instructed to tap as fast and as regularly as possible, lifting his or her finger at least 2cm above the table on each cycle. The force sensor surface was placed on the table and the patient is instructed to tap on the same spot every time. In middle finger speed tapping, the procedure was the same as with index finger speed tapping but switching to the middle finger.

The raw data files containing the force trace over time were preprocessed using a custom developed python script (we do not report the accelerometer data here). The script first discarded the first and last 0.5 sec of data from the recordings and then converted the force sensor trace into Newtons using a prior calibration table. It then smoothed the signal using a 160-sample Bartlett window (which amounted to approximately 53 msec at our sample rate). The script then detected the tap onset landmarks (a sudden impact) when the force exceeded .05 Newton. Tap offset (a release of contact from the tapping measurement surface) was detected when the force trace dropped below .05 Newton again at least 40 msec after the last tap onset. Similarly, the next onset was restricted to occur at least 75 msec after the last tap offset. All data files with their landmarks were then visually inspected and in a number of cases the .05 Newton onset/offset threshold was adjusted manually to compensate for the fact that some patients tapped too softly. We furthermore recorded the maximal tapping force between subsequent tap onset and tap offsets; the intervals between adjacent onsets, which will be referred to as the inter-tap-intervals (ITIs) in what follows; and the duration between the tap onset and tap offset (tap dwell phase duration). Next, we discarded the ITIs that were larger than 2000 msec since these marked pauses or interruptions in the patient's tapping behaviour (such as asking the experimenter whether to continue tapping) instead of the patient's motor control capacity. We also discarded ITIs shorter than 120 msec since there were disproportionately many as a result of double-tap recording.

### **3.2.e Mood test: Profile of Mood States**

Patients' mood was established using the Profile of Mood States (POMS) (Lorr, McNair, & Droppleman, 1971). The short form has 35 adjectives (items) of which the patient rates to what extent they are applicable to his or her mood over the last week, on a scale from 1 (not at all) to 5 (very strongly). The items load onto four category sub-scores: depression/anxiety, fatigue, vigor and hostility (Curran, Andrykowski, & Studts, 1995). We used the German translation (Bullinger, Heinisch, Ludwig, & Geier, 1990). This questionnaire was measured at PRE and POST. The experimenter reads each of the items to the patient, who then responds verbally. This is filled out on a sheet and at a later stage entered into a table independently by two researchers. The tables were then compared and discrepancies corrected by checking again the paper forms.

### **3.2.f Mood test: faces scale**

In order to obtain a quick estimate of the patient's mood throughout the therapy, we used a mood scale of faces (Andrews & Crandall, 1976; Kunin, 1955; McDowell, 2006). Patients were presented a list of smiley faces, ranging from very happy to very sad. Patients were asked at the PRE measurement session which face best represented how he or she is feeling by pointing to the corresponding face. At the beginning of each therapy session, patients were asked to point how they had felt since the previous session. At the end of each therapy session, too, patients were asked again how he or she has felt during the therapy session. Also, at the end of each joint therapy session, each patient was asked individually how he or she felt about the other patient with whom they received therapy. The patients pointed to one of the faces in such a way that this was not seen by the other patient, so as to avoid social pressure effects. Finally, during the POST measurement patients were again asked to select the face representing how he or she feels at present. The therapist wrote down the letter code corresponding to the chosen face without allowing the patient to see the letter in question.

### **3.2.g Ethics**

This study was performed in accordance with ethical guidelines proposed by the

Medical University Hannover (MHH). The protocol was approved by the ethics board on 20 April 2011 (nr. 1056-2011).

### 3.2.h Data analysis

We performed parametric ANOVA whenever the data quantity and distribution could reasonably be assumed to fulfil its assumptions. We detected deviations from sphericity using Mauchly's Test and whenever it was significant we applied the Greenhouse-Geisser correction. In those cases, we indicated significance as  $p_{GG}$  and omitted the uncorrected p-value for brevity. We report generalised effect sizes  $\eta_G^2$  (Bakeman, 2005). Groups were then compared using Tukey HSD contrasts.

## 3.3 Results

### 3.3.a Nine-hole pegboard test

We performed an ANOVA with time to complete the pegboard test as dependent variable and factors group (in-turn or together) and measurement (PRE or POST). Both groups performed the 9HPT faster after therapy [ $F(1,26)=21.35$ ,  $p<.0001$ ,  $\eta_G^2=.02$ ], but the in-turn group tended to improve more than the together group [ $F(1,26)=.70$ ,  $p=.065$ ,  $\eta_G^2=.004$ ] (Figure 1). There was no main effect of group [ $F(1,26)=.74$ ,  $p=.40$ ]. Some caution may be needed in interpreting this interaction effect, since the in-turn group performed the test slightly slower at the PRE measurement ( $M=72.4$  sec,  $SD=32.8$  sec) than the together group ( $M=42.5$  sec,  $SD=36.8$  sec). However, this difference was not significant [ $t(26.7)=1.13$ ,  $p=.27$ ]. Furthermore, the effect might reflect two patients with larger improvement/deterioration scores. However, these patients were not more than 3 SD away from the overall mean improvement or the mean improvement per group and were therefore not be discarded as outliers.

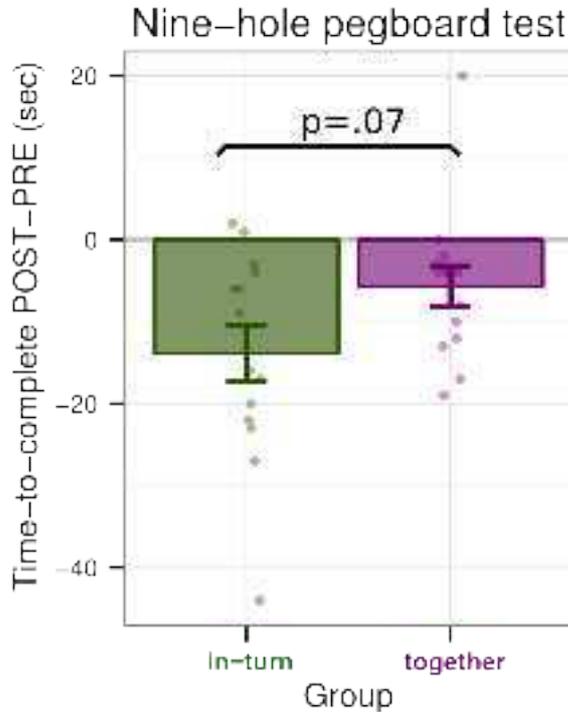


Figure 1: Nine-hole pegboard test measured as the difference in time-to-complete PRE-POST in seconds. A positive difference means that patients completed the Nine-hole pegboard test faster after therapy (POST).

### 3.3.b Finger tapping tests

#### i. Index finger unpaced tapping

The PRE measurement of one patient (in the in-turn group) was invalid due to technical reasons and this patient was therefore removed from further analysis. We pooled the taps that were recorded before and after each session and then computed the speed and variability as follows. Speed was calculated as the median of the intervals (in msec). Variability was calculated by first discarding the taps that were 3 SD longer or shorter than the mean for that block, taking the standard deviation of the remaining intervals and then divided it by the mean for that block to obtain the coefficient-of-variation (CV in percent).

There were no initial differences in tapping speed between the groups [ $t(18.8)=1.09$ ,  $p=.29$ ] or in tapping variability [ $t(24.8)=-1.2$ ,  $p=.24$ ].

We performed an ANOVA on the log-transformed median tapping interval with factors session (PRE, POST and the ten therapy sessions) and group (in-turn, together). The main effect of group was not significant [ $F(1,25)=.15$ ,  $p=.70$ ]. However, the main effect of measurement session was significant but became only a statistical trend after sphericity corrections [ $F(11,275)=2.20$ ,  $p=.01$ ,  $p_{GG}=.10$ ],  $\eta_G^2=.02$ ]. There was no interaction between group and session [ $F(11,275)=.72$ ,  $p=.72$ ,  $p_{GG}=.53$ ] (Figure 2).

We performed the same ANOVA with log-transformed coefficient-of-variability (CV) as dependent measure. We found no effect of group [ $F(1,25)=.09$ ,  $p=.76$ ], measurement session [ $F(11,275)=1.43$ ,  $p=.16$ ,  $p_{GG}=.20$ ] and no interaction [ $F(11,275)=1.51$ ,  $p=.12$ ,  $p_{GG}=.17$ ] (Figure 2).

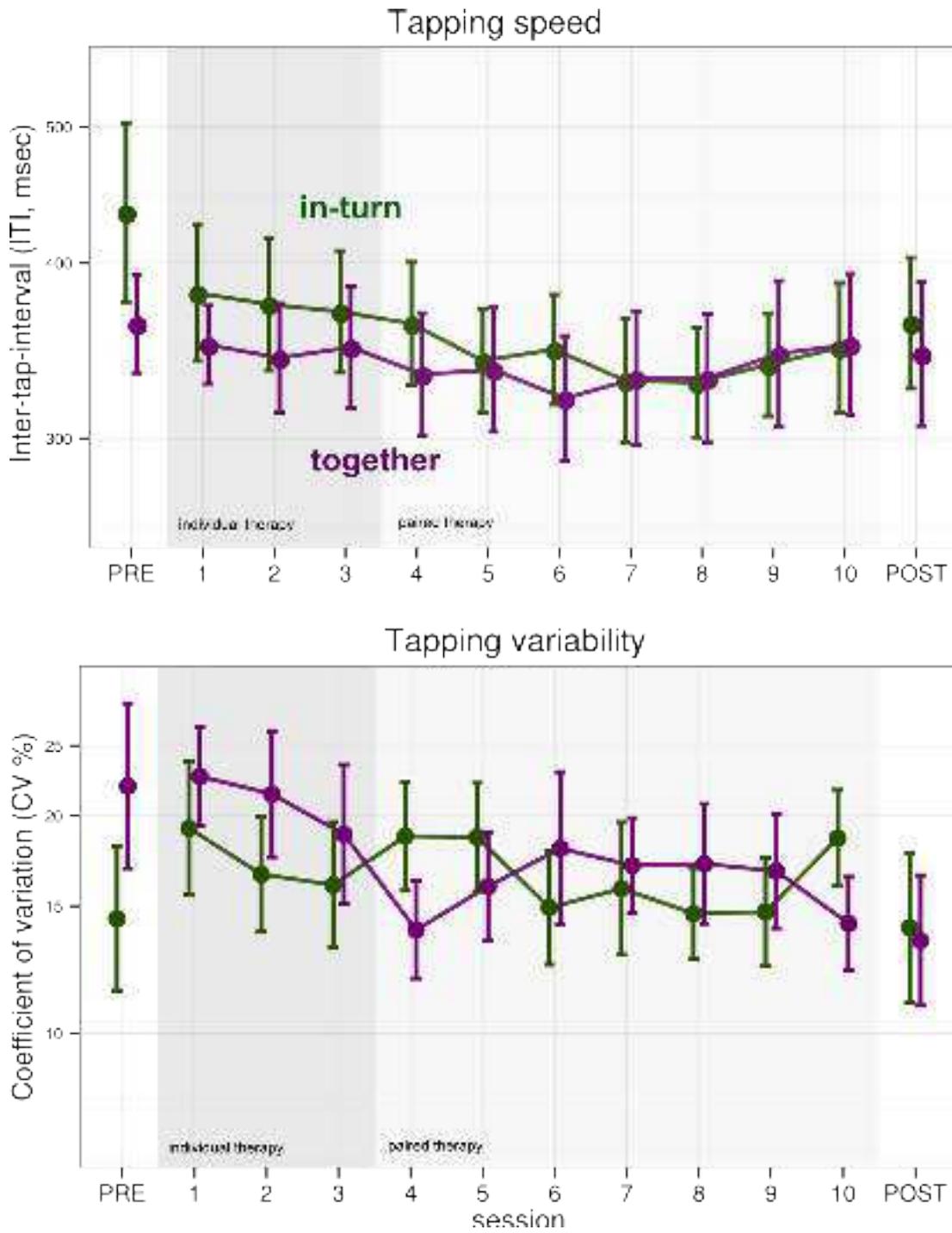


Figure 2: Index finger tapping speed and regularity. Error bars indicate the standard error of the mean.

**ii. Middle finger unpaced tapping**

Middle finger tapping was measured before (PRE) and after (POST) therapy. There were no differences in initial tapping speed [ $t(13.0)=1.76$ ,  $p=.11$ ] but there was a statistical trend for the in-turn group to be tapping more regular ( $M=17.2$ ,  $SD=9.9$  %CV) at the PRE measurement than the together group ( $M=30.2$ ,  $SD=19.0$  %CV) [ $t(24.0)=-1.77$ ,  $p=.09$ ].

An ANOVA with factors group (in-turn, together) and measurement session (PRE, POST) revealed no effect of group on middle finger tapping speed [ $F(1,24)=1.89$ ,  $p=.18$ ]. There was no effect of measurement session [ $F(1,24)=1.86$ ,  $p=.18$ ] and no interaction [ $F(1,24)=.20$ ,  $p=.66$ ].

The same ANOVA was performed with tapping variability as dependent variable. We found no effect of group [ $F(1,24)=.88$ ,  $p=.36$ ], recording session (PRE,POST) [ $F(1,24)=.85$ ,  $p=.36$ ]. There was a statistical trend for an interaction between group and session [ $F(1,24)=3.29$ ,  $p=.08$ ,  $\eta_G^2=.04$ ]. However, in light of the subtle differences in middle finger tapping variability at the PRE measurement, we interpret these findings as regression toward the mean.

**iii. Index-to-thumb paced tapping**

Two patients were eliminated from further analysis because during one session their tapping was too soft to be reliably assessed (both from the *together* group). We used circular statistics to assess the time-lock (synchronisation) between patients' finger tap onsets and the metronome click onsets. We then performed a repeated-measures ANOVA with factors group (together, in-turn) and measurement time point (PRE, INTER and POST) (Figure 3). The main effect of group was not significant [ $F(1,24)=2.54$ ,  $p=.12$ ]. There was a main effect of recording time-point [ $F(2,48)=10.98$ ,  $p_{GG}=.0001$ ,  $\eta_G^2=.09$ ]. This effect reflected the fact that patients' tapping was more synchronised after therapy relative to before [ $p=.036$ ]. There were no differences between the PRE and INTER measurements [ $p=.60$ ] or the POST and INTER measurements [ $p=.27$ ]. There was no interaction between group and time-point [ $F(2,48)=1.28$ ,  $p_{GG}=.29$ ].

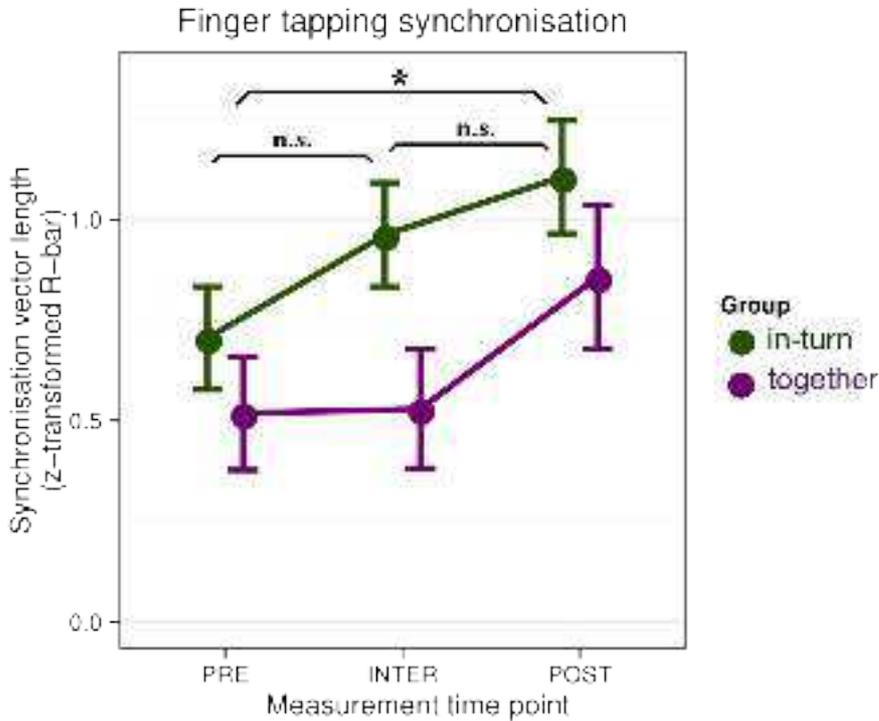


Figure 3: Synchronisation tapping performance before therapy (PRE), in between the individual and joint sessions (INTER) and after therapy (POST). Error bars indicate the standard error of the mean.

### 3.3.c Mood tests

For each factor of the Profile of Mood States (depression/anxiety, fatigue, hostility and vigor) we performed an ANOVA with factors group (together or in-turn) and timepoint (PRE, POST). We found that groups showed equal reductions in depression [ $F(1,26)=11.76$ ,  $p=.002$ ,  $\eta^2=.09$ ] and fatigue [ $F(1,26)=6.56$ ,  $p=.02$ ,  $\eta^2=.07$ ], whereas no change was found in vigor [ $F(1,26)=1.01$ ,  $p=.32$ ] although perhaps a trend for improvement in hostility [ $F(1,26)=3.95$ ,  $p=.06$ ]. There were no main effects of group [all  $F(1,26)<0.45$   $p>.51$ ] or interactions between group and timepoint [all  $F(1,26)<.19$ ,  $p>.67$ ] (Figure 4).

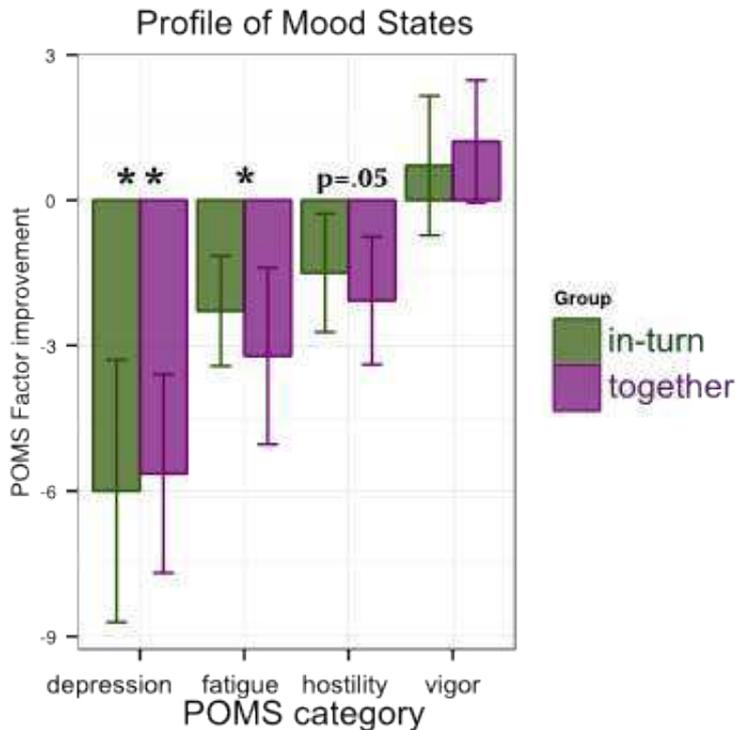


Figure 4: Results of the mood tests: patients show reductions in depression/anxiety and fatigue that are similar between groups (\* $p < .05$ , \*\* $p < .01$ ). A trend for improvement (decrease) in hostility is observed. Error bars indicate the standard error of the mean.

### 3.3.d Faces scale mood ratings

Patients were also asked to rate their own mood on the faces scale, both during the PRE and POST measurement sessions and during the therapy sessions. There were no differences in rating between the groups at the PRE measurement [Mann-Whitney U,  $Z = -1.22$ ,  $p = .22$ ]. The in-turn group's self-mood ratings improved during therapy [Friedman test  $\chi^2(11) = 27.36$ ,  $p = .004$ ] as did those of the together group [Friedman test  $\chi^2(11) = 36.08$ ,  $p = .0001$ ]. There were no differences in rating during the POST measurement session [Mann-Whitney U,  $Z = -.62$ ,  $p = .54$ ].

Patients were furthermore invited to rate how they experienced the therapy sessions. There was a tendency for the in-turn group to rate the first (individual) session

more positive than the together group [Mann-Whitney U,  $Z=-1.76$ ,  $p=.08$ ], although that first they still received therapy individually, but this difference had disappeared by the third session [Mann-Whitney U,  $Z=0$ ,  $p=1.00$ ]. During the paired therapy (sessions 4-10), the in-turn group became more positive as therapy progressed [Friedman  $\chi^2(6)=13.87$ ,  $p=.03$ ] but the together group stayed at the same level [Friedman  $\chi^2(6)=7.56$ ,  $p=.27$ ]. There were nevertheless no differences in rating between the groups at the last (tenth) session [Mann-Whitney U,  $Z=-.25$ ,  $p=.80$ ].

In the partner sympathy ratings, there were no differences in rating between the two groups in the first paired session (session 4) [Mann-Whitney U,  $Z=-.55$ ,  $p=.58$ ]. The in-turn group, showed a marked improvement in ratings of their therapy partner [Friedman  $\chi^2(6)=25.12$ ,  $p=.0003$ ] whereas the together group showed no change in rating [Friedman  $\chi^2(6)=4.98$ ,  $p=.55$ ] (Figure 5).

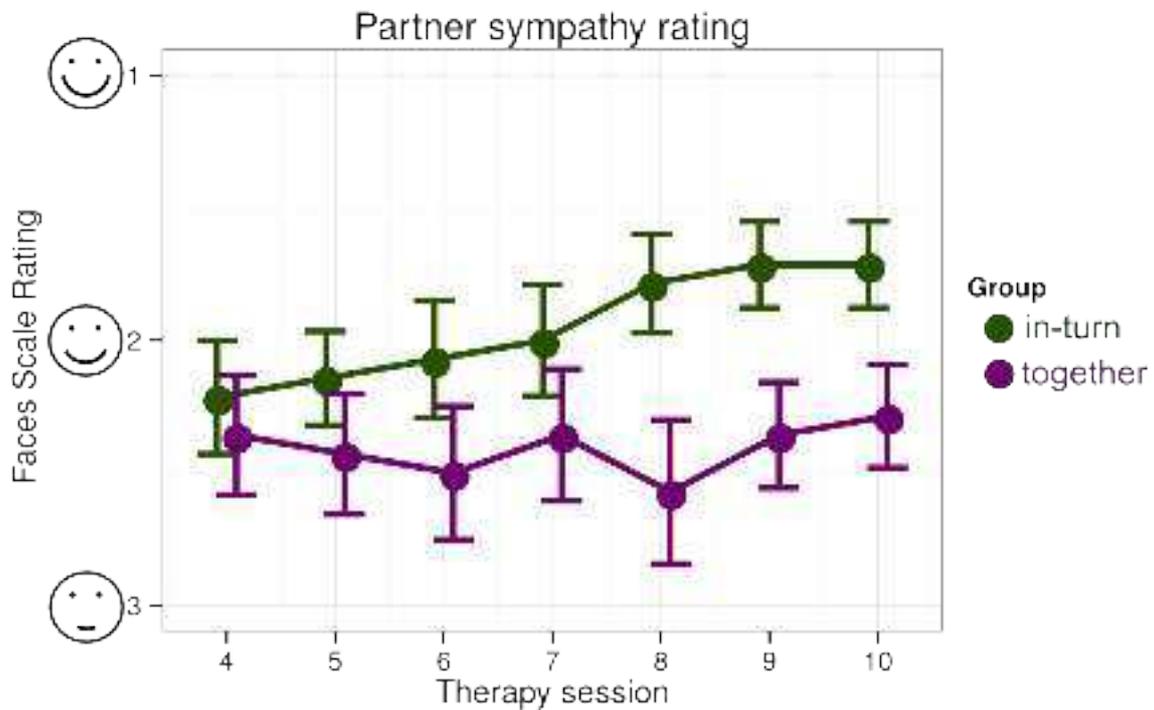


Figure 5: Patient's likeability rating of their therapy partner. We indicate mean and standard errors of the ratings for each group.

### 3.4 Conclusions

Our study was the first to implement music-supported therapy with pairs of patients instead of individually. We hypothesised that playing in synchrony would improve patient's social engagement and, through this greater engagement, improve their motor outcome. We controlled for potential benefits of patients sharing their musical rehabilitation experience (Overy, 2012) by having all patients receive therapy in pairs.

Firstly, our results reveal that music-supported stroke rehabilitation can be effective not only when patients are treated individually (as in previous studies) but also in pairs. We found clear improvements in patients' fine motor control in the nine-hole pegboard test and

synchronisation tapping. The finding that music-supported rehabilitation is effective in pairs has practical implications. Paired therapy could considerably reduce the time investment on the part of the treatment personnel.

Surprisingly, we found no clear improvements in index or middle finger tapping. This is in contrast to previous studies of music-supported therapy who found improvements in finger tapping frequency (Amengual et al., 2013; Schneider et al., 2007). In the case of finger tapping speed, the absence of overall improvement could be due to the fact that finger tapping speed showed a u-shaped curve (Figure 2). Patients appeared to be improving in the first half of the therapy but then showed a tendency for a rebound in the second half of the therapy. Alternatively, there may be a considerable effect of which therapist implements music-supported therapy on the therapy outcome.

The nine-hole pegboard test showed a difference in rehabilitation outcome between patients playing in synchrony and patients playing in turn. Contrary to our hypothesis, which was that patients in the together group would show the greatest benefit, this test showed that patients in the in-turn group benefited most.

The fact that both groups performed mostly similar is an indication that playing in synchrony might not lead to benefits in rehabilitation. We speculate that patients in the former group may benefit from the opportunity to learn from observation. In healthy participants, seeing others perform a motor task leads to motor learning in the observer (Hodges, Williams, Hayes, & Breslin, 2007; McCullagh, Weiss, & Ross, 1989; Wulf & Mornell, 2008). Indeed, action observation has been proposed recently as a tool for motor-rehabilitation after stroke (Garrison, Aziz-Zadeh, Wong, Liew, & Winstein, 2013; Garrison, Winstein, & Aziz-Zadeh, 2010; Sale & Franceschini, 2012). In particular, observers appear to benefit from observing both experts and novices perform a motor task, thus learning from errors as well as exemplary performance (Andrieux & Proteau, 2013). This finding suggests that stroke patients undergoing rehabilitation may benefit from observing a therapist perform movements as well as a patient peer, as they did in our in-turn group. In this way, music-supported therapy might optimise the patients' therapeutic benefit.

Alternatively, it is possible that the simultaneously occurring sounds in the together condition confused patients, preventing them from dissociating sounds that they

self-generated from those that were generated by their partner. This could have prevented the motor system from learning from its auditory feedback (Altenmüller et al., 2009). A future study could remedy this problem by providing the two patients in each pair separate headphones in which their own sounds are louder than those of their partner.

Furthermore, results indicate that patients in the in-turn group grew to like each other more over the course of therapy. This is contrary to previous findings where people moving in synchrony liked each other more (Hove & Risen, 2009) than those who did not. Perhaps this difference is due to impaired auditory-motor functioning in stroke patients (Rodriguez-Fornells et al., 2012). We found no differences in finger synchronisation tapping performance between the groups, suggesting that the overall improvement in synchronisation was due to a general improvement in motor capacity and not the fact that one group trained synchronising during therapy. Therefore, it may be so demanding to patients to synchronise their playing to their partner that the mechanisms that usually mediate synchrony-induced social effects (Wiltermuth, 2012; Wiltermuth & Heath, 2009) were suppressed.

A limitation of this study is that we have not tested a control group who did not receive any musical intervention. As a result, the effects that do not differ between groups cannot strictly be attributed to the musical intervention. The advantage, however, of this approach is that any differences between the groups are likely due to the principal experimental manipulation (playing together vs. playing in turn). Our patient sample was relatively small and heterogeneous and the exact lesion sites of the stroke were unknown to us. Future studies could correlate lesion localisation maps to performance and functional motor outcome of patients undergoing music-supported therapy in order to establish which patient groups might benefit maximally from music-supported therapy. Finally, we have not been able to synchronise our recordings from the two keyboards used in the joint therapy sessions. As a result, we could not compare synchronisation performance within therapy.

## References

- Altenmüller, E., Marco-Pallares, J., Münte, T. F., & Schneider, S. (2009). Neural reorganization underlies improvement in stroke-induced motor dysfunction by music-supported therapy. *Annals of the New York Academy of Sciences*, 1169, 395–405. doi:10.1111/j.1749-6632.2009.04580.x

- Amengual, J. L., Rojo, N., Veciana de las Heras, M., Marco-Pallarés, J., Grau-Sánchez, J., Schneider, S., ... Rodríguez-Fornells, A. (2013). Sensorimotor Plasticity after Music-Supported Therapy in Chronic Stroke Patients Revealed by Transcranial Magnetic Stimulation. *PLoS ONE*, *8*(4), e61883. doi:10.1371/journal.pone.0061883
- Andrews, F. M., & Crandall, R. (1976). The validity of measures of self-reported well-being. *Social Indicators Research*, *3*(1), 1–19. doi:10.1007/BF00286161
- Andrieux, M., & Proteau, L. (2013). Observation learning of a motor task: who and when? *Experimental Brain Research*, *229*(1), 125–137. doi:10.1007/s00221-013-3598-x
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods*, *37*(3), 379–384.
- Caeiro, L., Ferro, J. M., & Costa, J. (2013). Apathy secondary to stroke: a systematic review and meta-analysis. *Cerebrovascular diseases (Basel, Switzerland)*, *35*(1), 23–39. doi:10.1159/000346076
- Calautti, C., Jones, P. S., Persaud, N., Guincestre, J.-Y., Naccarato, M., Warburton, E. A., & Baron, J.-C. (2006). Quantification of index tapping regularity after stroke with tri-axial accelerometry. *Brain Research Bulletin*, *70*(1), 1–7. doi:10.1016/j.brainresbull.2005.11.001
- Dimyan, M. A., & Cohen, L. G. (2011). Neuroplasticity in the context of motor rehabilitation after stroke. *Nat Rev Neurol*, *7*(2), 76–85. doi:10.1038/nrneurol.2010.200
- Garrison, K. A., Aziz-Zadeh, L., Wong, S. W., Liew, S.-L., & Winstein, C. J. (2013). Modulating the Motor System by Action Observation After Stroke. *Stroke; a journal of cerebral circulation*. doi:10.1161/STROKEAHA.113.001105
- Garrison, K. A., Winstein, C. J., & Aziz-Zadeh, L. (2010). The Mirror Neuron System: A Neural Substrate for Methods in Stroke Rehabilitation. *Neurorehabilitation and Neural Repair*, *24*(5), 404–412. doi:10.1177/1545968309354536
- Glass, T. A., Matchar, D. B., Belyea, M., & Feussner, J. R. (1993). Impact of social support on outcome in first stroke. *Stroke*, *24*(1), 64–70. doi:10.1161/01.STR.24.1.64
- Grau-Sánchez, J., Amengual, J. L., Rojo, N., Heras, M. V. de las, Rubio, F., Altenmüller, E., ... Rodríguez-Fornells, A. (2013). Plasticity in the sensorimotor cortex induced by Music-supported therapy in stroke patients: a TMS study. *Frontiers in Human Neuroscience*, *7*, 494. doi:10.3389/fnhum.2013.00494
- Heller, A., Wade, D. T., Wood, V. A., Sunderland, A., Hewer, R. L., & Ward, E. (1987). Arm function after stroke: measurement and recovery over the first three months. *Journal of Neurology, Neurosurgery, and Psychiatry*, *50*(6), 714–719.
- Hodges, N. J., Williams, A. M., Hayes, S. J., & Breslin, G. (2007). What is modelled during observational learning? *Journal of sports sciences*, *25*(5), 531–545. doi:10.1080/02640410600946860
- Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition*, *27*(6), 949–960. doi:10.1521/soco.2009.27.6.949
- Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized Drumming Enhances Activity in the Caudate and Facilitates Prosocial Commitment - If the Rhythm Comes Easily. *PLoS ONE*, *6*(11), e27272. doi:10.1371/journal.pone.0027272
- Kunin, T. (1955). The Construction of a New Type of Attitude Measure. *Personnel Psychology*, *8*, 65–77. doi:10.1111/j.1744-6570.1955.tb01189.x
- Launay, J., Dean, R. T., & Bailes, F. (2013). Synchronization can influence trust following virtual interaction. *Experimental psychology*, *60*(1), 53–63. doi:10.1027/1618-3169/a000173
- Lorr, M., McNair, D., & Droppleman, L. (1971). Manual: profile of mood states. *San Diego, CA: Educational and Industrial Testing Service*.
- Mathiowetz, V., Volland, G., Kashman, N., & Weber, K. (1985). Adult norms for the Box and Block Test of manual dexterity. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *39*(6), 386–391.
- McCullagh, P., Weiss, M. R., & Ross, D. (1989). Modeling considerations in motor skill acquisition and performance: an integrated approach. *Exercise and sport sciences reviews*, *17*, 475–513.

- McDowell, I. (2006). *Measuring health: a guide to rating scales and questionnaires*. Oxford University Press.
- Narme, P., Tonini, A., Khatir, F., Schiaratura, L., Clément, S., & Samson, S. (2012). Non pharmacological treatment for Alzheimer's disease: comparison between musical and non-musical interventions. *Geriatric et psychologie neuropsychiatrie du vieillissement*, 10(2), 215–224. doi:10.1684/pnv.2012.0343
- Okuno, R., Yokoe, M., Akazawa, K., Abe, K., & Sakoda, S. (2006). Finger taps movement acceleration measurement system for quantitative diagnosis of Parkinson's disease. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, Suppl*, 6623–6626. doi:10.1109/IEMBS.2006.260904
- Overy, K. (2012). Making music in a group: synchronization and shared experience. *Annals of the New York Academy of Sciences*, 1252(1), 65–68. doi:10.1111/j.1749-6632.2012.06530.x
- Parker, V. M., Wade, D. T., & Langton Hewer, R. (1986). Loss of arm function after stroke: measurement, frequency, and recovery. *International Rehabilitation Medicine*, 8(2), 69–73.
- Rodriguez-Fornells, A., Rojo, N., Amengual, J. L., Ripollés, P., Altenmüller, E., & Münte, T. F. (2012). The involvement of audio-motor coupling in the music-supported therapy applied to stroke patients. *Annals of the New York Academy of Sciences*, 1252(1), 282–293. doi:10.1111/j.1749-6632.2011.06425.x
- Sale, P., & Franceschini, M. (2012). Action observation and mirror neuron network: a tool for motor stroke rehabilitation. *European journal of physical and rehabilitation medicine*, 48(2), 313–318.
- Sandin, K. J., Cifu, D. X., & Noll, S. F. (1994). Stroke rehabilitation. 4. Psychologic and social implications. *Archives of physical medicine and rehabilitation*, 75(5 Spec No), S52–55.
- Schneider, S., Schönle, P. W., Altenmüller, E., & Münte, T. F. (2007). Using musical instruments to improve motor skill recovery following a stroke. *Journal of Neurology*, 254(10), 1339–1346. doi:10.1007/s00415-006-0523-2
- Valdesolo, P., & Desteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion (Washington, D.C.)*, 11(2), 262–266. doi:10.1037/a0021302
- Ward, N. S., & Cohen, L. G. (2004). Mechanisms underlying recovery of motor function after stroke. *Archives of Neurology*, 61(12), 1844–1848. doi:10.1001/archneur.61.12.1844
- Wiltermuth, S. S. (2012). Synchrony and destructive obedience. *Social Influence*, 7(2), 78–89. doi:10.1080/15534510.2012.658653
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*, 20(1), 1–5. doi:10.1111/j.1467-9280.2008.02253.x
- Wulf, G., & Mornell, A. (2008). Insights about practice from the perspective of motor learning: a review. *Music Perform Res*, 2, 1–25.

## Acknowledgements

This work was supported by the EBRAMUS, European Brain and Music Ph.D. Grant to FTV (ITN MC FP7, GA 238157). We are indebted to Britta Westner, M.A., for running a pilot study. We furthermore would like to thank Mr. Richter for valuable feedback about music-supported therapy. Also, Sabine Schneider kindly shared her knowledge of the previous implementation of the music-supported therapy that formed the basis for the current therapy program. We thank all the medical staff (doctors and nurses) in the Hessisch Oldendorf clinic for their cooperation and for indicating to us patients who might

be suitable for inclusion in the study. Finally, but not least, we wish to extend our gratitude to all the patients who devoted their time to participation in this study. We hope the therapy may benefit them in their daily lives.

## III.4. Discussion

---

The present part of the thesis investigated the potential of auditory feedback in motor rehabilitation after stroke (study III.2) as well as the potential of synchronisation of this auditory feedback between patients (study III.3).

Auditory-motor integration has been proposed to be a crucial ingredient for music-supported therapy. A patient was found to show motor co-activation during passive music listening after therapy (Rodriguez-Fornells et al., 2012; Rojo et al., 2011). However, we asked whether this motor activation was *critical*, i.e. causally involved in motor improvement, or merely a side-effect, i.e. associatively involved (see part I of this thesis).

### 4.1 Auditory feedback

Among the key findings reported here are that disruptions in timing of auditory feedback do not disrupt motor rehabilitation. On the contrary, disruptions in auditory feedback may enhance the outcome along several clinical measures (study III.2). What are the implications of this finding for the theoretical underpinning of auditory-motor involvement in motor stroke rehabilitation?

First of all, the finding that learning is not disrupted by disruptions in feedback makes it less likely that patients use sound as a feedback signal (Shadmehr & Krakauer, 2008). Instead, they may have been using other feedback signals such as proprioception. In section 1.2 we have shown that only limited evidence exists that proprioception is damaged in stroke patients. However, it is possible that the temporal noise we have added to the sounds is so small relative to the noise in a patient's (damaged) auditory-motor system that the effect of temporal jitter is negligible.

We will return to the question of auditory feedback in the general discussion in section V.1.

## 4.2 A note about mood improvements

Does music-supported therapy improve patients' mood? To date, no study has evaluated the effect of music playing on the mood of stroke patients. Studies III.2 and III.3 of this thesis found a similar pattern of mood improvements in the two independent patient groups. A standard measure of mood (*Profile of mood states*) revealed reductions in depression/anxiety as well as fatigue (studies III.2 and III.3) as well as, perhaps, a reduction in hostility (study III.3). These mood improvements did not vary according to our experimental manipulation, such as auditory feedback or whether patients played together or alone. Whether this mood improvement is a direct result of music therapy is not clear at this point. The reason is that in both studies we contrasted different implementations of music therapy instead of testing them against a control group. In this way we provided a better controlled setting for the testing of the specific effects we were interested in (auditory feedback or synchronisation), but the downside is that we cannot provide definite conclusions concerning quantities that we found not to be different between groups. For example, the reduction in depression/anxiety and fatigue may be a general effect of the therapeutic setting in which patients were immersed (the rehabilitation hospital) or a side-effect of their motor improvement. The only study to date which compared music-supported training with a control group (Schneider, Münte, Rodriguez-Fornells, Sailer, & Altenmüller, 2010) did not include mood measurements. However, the fact that our results are similar across the two (independent) studies provides some support for the idea that the mood improvements we found are a result of music-supported therapy. In addition, mood improvements have often been associated with exposure to music (Thompson, Schellenberg, & Husain, 2001) and in particular when the music is familiar (Pereira et al., 2011).

## 4.3 Implications for clinical practice

The results presented in this part of the thesis have a number of clinical

consequences. For the first time, we investigated the time course of the motor skill improvements during therapy. Our crucial finding was that the majority of the improvement in speed happens in the first few therapy sessions (study III.2).

A second finding was that therapy in pairs appears equally effectively as therapy alone (comparing studies III.2 and III.3). This suggests that music-supported therapy may well be offered in pairs. The fact that therapy in which patients play in turn, observing and pausing while the other person plays (study III.3), is as effective (if not more effective) than therapy where patients play together implies that it will be easy for therapists to divide their time between the two patients.

Do the present findings imply that we should recommend clinical practitioners use a delayed feedback piano? We argue against prematurely applying these findings in clinical settings. First of all, the effects are relatively subtle and might disappear at particular times in the therapy (such as the finger tapping regularity difference in study III.2 vanishing at the last session). Therefore the gain in motor performance is relatively small. Secondly, due to the limited size of our experimental groups, it remains difficult to rule out initial differences between them. Furthermore, we found that the auditory-motor delay detection thresholds are actually *higher* in the group that had been practising with a delayed feedback piano (study III.2). At this point, it is unclear what other (possibly undesirable) consequences such higher thresholds have for patient functioning. It is unlikely that these will cause negative behavioural effects, since many studies routinely manipulate auditory-motor (or multisensory) synchrony in healthy participants (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Keetels & Vroomen, 2012; Sugano, Keetels, & Vroomen, 2010) and did not report any behavioural side-effects. However, we would like to call for detailed investigation of such side-effects in future studies.

## **IV. Over-learning in Musical Experts**

## IV.1. Introduction

---

Expert musicians have spent many years practising movements with the goal of producing a desired auditory output. In this part of the thesis, we ask how auditory feedback is used by expert musical performers. To answer this question, we will develop novel methods of analysis that investigate fine properties of timing control in these experts. The task we investigate is scale playing, as it provides a close parallel to the regularity learning used in parts II and III of this thesis.

### 1.1 Choice of experimental paradigm

In order to be able to study auditory-motor crosstalk by comparing three different populations (healthy nonmusicians, stroke patients, and musical experts) the question of what kind of task we will investigate naturally arises. In the case of musical experts discussed in this part of the thesis, we have chosen to turn to the playing of musical scales for a variety of reasons.

Firstly, scale playing is a technical task that does not involve musical interpretation to the extent that playing musical compositions does. Playing musical compositions engages the musician in an artistic process during which he or she interprets the musical structure and uses timing cues (among other things) to signal landmarks in this structure (Palmer, 1997). Studying how musical experts perform the task of playing musical materials, therefore, would introduce an insurmountable interpretation bias: any control differences arising between the three groups could just as easily be due to the fact that musicians are doing something fundamentally different from stroke patients or healthy non-musicians. Furthermore, differences between pianists could arise from their differential interpretation of the music instead of differences in motor control.

Secondly, musicians have practiced scale playing for large amounts of time, as this

skill is essential to the repertoire of any musician (though non-classical pianists may differ from, for example, jazz pianists in this respect). This allows us to avoid the problem that a piece is practiced a great deal by some experimental participants, but not at all by others.

Fourthly, musical scale playing has previously been used to measure motor control. For example, it has been proposed as a measure of severity for musician's dystonia, a movement disorder (see study VIII.2 in this thesis).

Finally, playing musical scales is a task that even very advanced musicians can take seriously and will not discard as an artificial exercise done only to do the experimenter a favour.

However, a downside to the choice of musical scales is that the mechanisms underlying timing in musical scales are poorly understood. As a result, the studies we report in this part of the thesis will have a strong methodological emphasis. We develop the analytical machinery that enables us to understand the relation between motor performance and auditory feedback in experts playing scales.

## **1.2 Piano timing measurements: methodological considerations**

Among studies investigating musical performance, piano playing has received the most attention (Furuya & Altenmüller, 2013; Jabusch, 2006). This is not because it is necessarily the most interesting instrument, but rather because of the relative simplicity of the measurement of piano movements. Sounds are generated in a piano by a swift pressing down of the key which is mechanically connected to a hammer. The hammer hits a string which as a result starts to vibrate at a frequency that is determined by its length and tension.

The first systematic investigation of timing in piano performance is due to Seashore (1938). Seashore sought to describe the mechanical properties of the piano and the way these translate into sounds, proposing the following (p.227):

“It makes no difference whether the key is struck by an accelerating, retarding, even, or any form of irregular movement; the only significant thing the player controls in the stroke is the velocity of the key at the exact moment that it throws off the hammer.”

In other words, the sound generated by the piano is determined only by which notes are struck when and with what keystroke velocity. For any note, only three parameters would have to be measured to be able to reproduce a musical performance: (I) the string (note), (II) the hammer-string contact time, (III) the keystroke velocity. This observation formed the basis of the later development of the MIDI (Musical Instrument Digital Interface). MIDI is the protocol used for measurements throughout this part of the thesis by which musical instruments can communicate digitally to one another. For example, the interface allows a digital musical keyboard to communicate with a personal computer or other digital musical instruments. The protocol essentially consist in communicating the three parameters for each keystroke in binary form.<sup>4</sup>

In order to measure keystroke onset, offset and velocity, Seashore used optical cameras (the *Iowa piano camera*) mounted inside the piano. Each camera was able to record the onset or offset of the movement of each hammer. This output, essentially a *piano roll*, enabled the researchers to determine the timing of each note offline. Using these methods, Seashore was able to investigate pianists playing the same piece of music multiple time. The observation was that the timings were remarkably consistent across two or three performances of the same piece, thus validating that the timing measurement technique revealed a meaningful measure of performance. Indeed, it was shown later that keystroke timing and keystroke velocity can be varied independently by the pianist and that these two are used to signal the metric structure of the piece played (Sloboda, 1983). This finding was corroborated by the observation that the timing patterns of, for example, a Chopin Étude showed periodicities of roughly four bars (Shaffer, 1984a, 1984b). That is, timing was essentially used as a means to signal structural properties of the musical interpretation (see also Repp, 1992a, 1992b, 1999a). This idea is supported by the observation that more advanced pianists show *more* temporal deviation away from the musical score than beginners (Sloboda, 2000). Surprisingly, such deviations were reduced but still present when pianists were instructed to play metronomically and not expressively (Repp, 1999b). This suggests that pianists reduced their temporal deviations to below the perceptual threshold, but were unaware of maintaining traces of them.

How is the signalling of musical structure achieved? Typically, musical performers

---

<sup>4</sup> Recently, piano performance researchers disagree as to whether these three parameters are the only ones that determine the sound on a given piano (Furuya & Altenmüller, 2013)

slow down at the end of musical phrases (Repp, 1992a). This led to the suggestion that local slowing down signals structural boundaries (Penel & Drake, 1998) due to a particular grouping of the musical material (Lerdahl & Jackendoff, 1977; Todd, 1985). The phenomenon appears also in language (Turk & Shattuck-Hufnagel, 2007) and is reminiscent of timing curves in locomotion (Friberg & Sundberg, 1999).

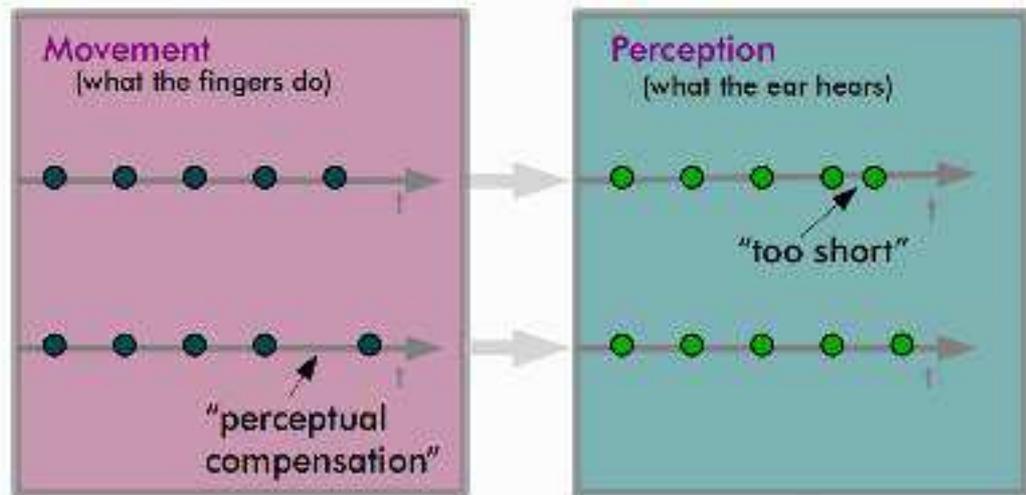
For a review of timing in piano performance, see Palmer (1997). For a recent review of piano measurements methodology, see Jabusch (2006).

### 1.3 Musical performance and the perception-action interface

The *perceptual hypothesis* provides an explanation for the interaction between perception and performance and is essential in studies IV.2 and IV.3 in this part of the thesis. The basis for this hypothesis is the finding that the loci of typical “expressive” lengthening are also places where perceptual accuracy is less (Repp, 1992a). That is, participants had more trouble detecting lengthening in precisely those notes that pianists typically prolong in expressive performance. The participants in the study were able to reliably pick out timing errors (though not always exactly locate them) of 56ms down to 23ms (6.7% of the interval).

This finding led to the formulation of the *perceptual hypothesis* (Drake, 1991, 1993; Penel & Drake, 1998). The proposal is that timing deviations in playing arise from the way we perceive the sounds. If, by hypothesis, some intervals sound shorter than others, the musical performer will compensate by playing those intervals longer (Illustration 7). In other words, the way people play is meant to counteract the distortions of temporal intervals as they are perceived.

The perceptual hypothesis is of interest in non-expressive and relatively “phrase-less” materials such as musical scales. The prediction is that consistent timing deviations are expected if the regularity of temporal intervals is not veridical, that is, if some intervals sound longer or shorter than others.



*Illustration 7:* Explanation of the perceptual hypothesis. By hypothesis, a systematic perceptual bias exists that makes certain intervals sound longer (or shorter) than others. In our example, the last interval sounds shorter than the previous intervals. If the musician's task is to produce as regular a scale as possible, she will play the final interval longer so that all intervals sound equally long.

## 1.4 Musical scales

This part of the thesis will investigate auditory-motor integration in expert musicians using scale playing as an example. In this section, we briefly introduce the movement and discuss previous studies that have investigated this task.

### 1.4.a Description of the scale playing movement

To play a two-octave scale, the pianist's task is to press fifteen adjacent keys serially (Illustration 8). Playing the ascending scale with the right hand, the thumb starts out on the c, then the index and middle finger press the adjacent d and e. Now the thumb has to move underneath the index and middle finger to strike the next key, f. The role of this manoeuvre is to move the position of the hand rightward. Next the index finger moves over the thumb to strike the g. The middle and ring finger can then in turn strike the a and b. At this point, another thumb-under movement is performed to strike the c, and then the entire movement

is repeated for the next octave, except for the last note (c) which is played with the little finger instead of with the thumb. The left hand movements are analogue for the descending scale starting with the thumb on c". In the studies we report here (studies IV.2, IV.3, IV.4, VIII.1 and VIII.2), participants played the scales with the conventional fingering (123123412312345 and reverse, where the fingers are numbered from 1=thumb to 5=little finger).



*Illustration 8:* Visualisation of the scale playing movement. Photos are ordered from left to right chronologically.

### 1.4.b Timing control in scale playing

The majority of studies into timing in piano playing investigated the playing of “ecological musical stimuli,” by which we mean musical compositions and not technical exercises such as scales. There are a few exceptions and now we turn to those studies that investigated timing control in playing musical scales.

Wagner (1971) sought to quantify playing precision of piano students. He instructed his participants to play piano scales over two octaves at a tempo that was indicated by a concurrent metronome. Using the piano roll, he was able to quantify the onset of each note and proceeded to calculate the intervals between subsequent notes within the scale. The standard deviation of these intervals is a measure of the regularity of playing: a larger standard deviation indicates more variability and thus less precision whereas a smaller standard deviation indicates higher precision. The main finding was that this standard deviation (which we will refer to as *unevenness*) does not increase linearly with tempo.

Rather, at very slow or very fast tempi, pianists were more uneven, and at intermediate tempi more even. That is, the unevenness was a U-shaped function of playing tempo, and pianists played most evenly at a tempo of around 8 keystrokes per second.

The question arose as to whether these deviations were the same across the notes in the scale. MacKenzie & Van Eerd (1990) replicated the finding by Wagner (1971) (except pianists optimal tempo was at slightly slower tempi than previously found) and also are the first to report the observation that note timing differs by position in the scale. Generally, the notes around the thumb-passage were played slower, indicating the greater difficulty involved in this movement.

How does scale playing precision relate to expertise? Jabusch, Yong, & Altenmüller (2007) investigated musical skill learning in 30 children between 8 and 17 years old. Scale playing unevenness correlated negatively with accumulated lifetime practice time and current daily practice time. Investigating more advanced students, a longitudinal study involving nineteen pianists at the end of their training at the Hanover conservatory was conducted (Jabusch, Alpers, Kopiez, Vauth, & Altenmüller, 2009). Their scale playing was measured at two time points, separated by about 27 months. Scale playing temporal unevenness at baseline correlated with the total accumulated life practice time, and the improvement in scale playing correlated with the accumulated practice time between the two measurements. No correlation was found with the age of commencement of playing.

Finally, we will point out that scale playing has also been used in clinical contexts, by employing scale playing unevenness properties as a diagnostic for *musician's dystonia*, which is a task-specific loss of voluntary muscular control of the hands (Altenmüller & Jabusch, 2010) (see also study VIII.2 in the present thesis for an example use of this metric). The authors (Jabusch, Vauth, & Altenmüller, 2004) developed a measure of scale playing unevenness that is a simple extension of the metric explained previously (Wagner, 1971). The procedure is as follows (see Illustration 9). For each (correct) scale run, the intervals between subsequent note onsets are determined. The standard deviation of these intervals is determined for each scale run and then the median is taken for each playing direction (ascending or descending). The principal finding was that dystonic patients exhibit a larger standard deviation of the inter-keystroke-intervals (sdIOI) than a control group. That is, their scale playing was more uneven in time than that of healthy controls. Further, the

playing unevenness correlated with dystonic severity as established by physicians inspecting the patients through video (Jabusch et al., 2004).

The neural correlates of temporal precision in scale playing remain largely elusive. The only study to date investigating this issue found that the playing unevenness correlated with grey matter distribution in the putamen (Granert, Peller, Jabusch, Altenmüller, & Siebner, 2011). Both normal and dystonic pianists that played more temporally precisely had a smaller regional middle putamen grey matter volume. This was true for the left hemisphere, and a trend was found for the same correlation in the right hemisphere. This finding suggests that subcortical structures may play a critical role in controlling scale playing timing.

# Interval-based Scale Analysis

Wagner (1971); Jabusch, Vauth, & Altenmüller (2004)

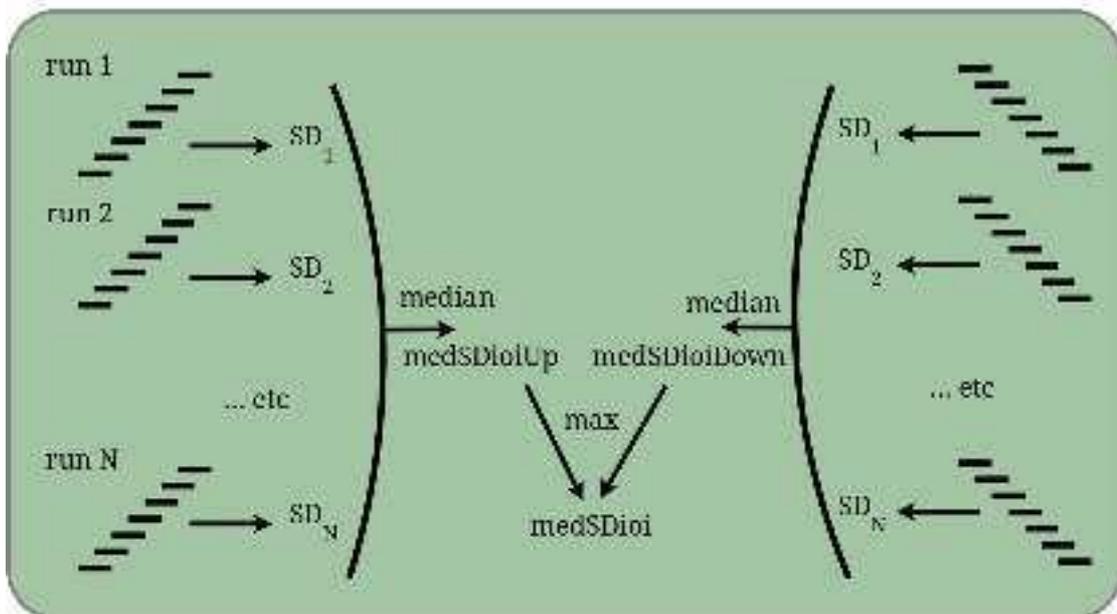
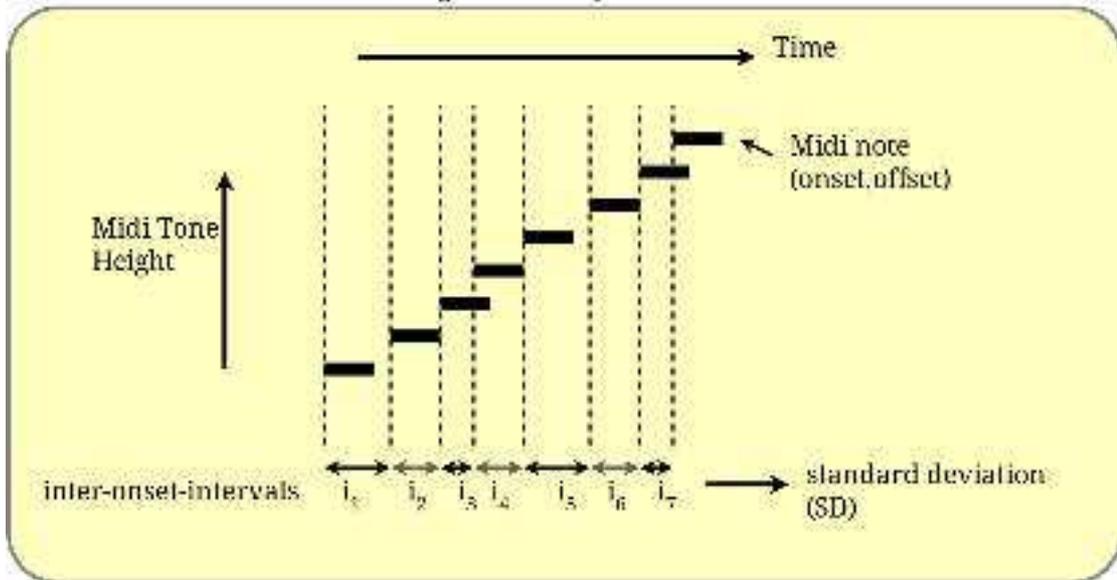


Illustration 9: Interval-based Scale Analysis.

## 1.4.c Towards a new model of timing in piano performance

The *unevenness* measure described in the previous section proceeds by pooling the intervals between the subsequent keystrokes in the scale. The tacit assumption is that the intervals between successive keystrokes are independently and identically distributed (IID).

Simply put, the assumption is that the intervals are, on average, equally long. (There are also other assumptions that will not concern us here.) However, this assumption is contradicted by the finding that keystrokes around the thumb-transitions tend to be longer than the others (MacKenzie & Van Eerd, 1990).

Secondly, the individual intervals do not always inform us about individual *note* timing deviations (Illustration 10). Imagine that one note is played slightly late. The two adjacent intervals will be affected (one being too long and the other being too short). As a result, an interval-based timing analysis will not be able to single out individual keystroke timings.

There is a need for a method of analysis that is able to single out timing values of individual keystrokes. Further, if the timings of the keystrokes in a scale are not identically distributed, it is necessary to disentangle consistent temporal deviations across trials from variability between trials. The studies presented in this chapter propose methods of analysis that address these two concerns. Using these tools, we can ask whether and how motor performance is influenced by auditory feedback.

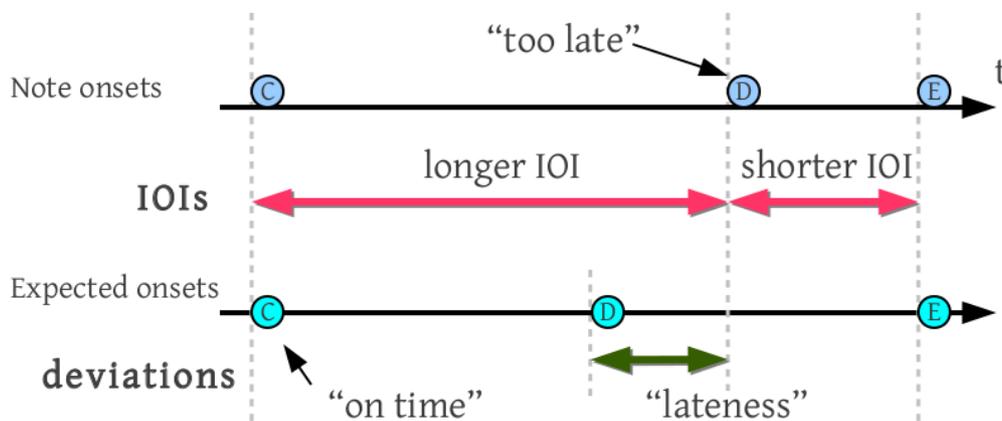


Illustration 10: Interval-based vs. note-based timing

## 1.5 Aims and hypotheses

The question guiding this part of the thesis is to what extent musicians use and have used auditory feedback in timing regularity. Musical experts have already accumulated thousands of hours of motor learning. In playing a scale, we hypothesise that some transitions between movements are more difficult than others. As a result, some transitions will take longer than others and consistently so. The task that musicians have been engaged in when practising scales is to make their movements more regular. If they use auditory feedback for this learning, this has the following important consequences. First, we expect that the irregularities in the timing will be regularised only until they are no longer audible. Secondly, the remaining irregularities should be very consistent and may vary considerably between pianists. The implications of this are important. It means that assumptions underlying the calculation of unevenness of the intervals are violated. Therefore, we will develop a novel analysis method that is hoped to capture this fact.

This part of the thesis has the following aims:

1. Quantify the amount of timing deviations that are consistent across trials and those that vary from trial to trial (**study IV.2**). The former will be referred to as irregularity and the latter as instability.
2. Test whether irregularity traces are below the perceptual threshold (**study IV.3**).
3. Establish the necessity of auditory feedback in formation of timing deviation (**study IV.3**).
4. Develop and test a model that captures consistent and inconsistent timing deviations from regularity in musical experts. In particular, we document and model timing deviations at various speeds (**study IV.4**), investigate (circadian) variations throughout the day (**study VIII.1**). We also show how timing deviations can be taken as a proxy for severity of musicians' dystonia, and how alterations of (somatosensory) feedback as provided in retraining therapy may benefit patients (**study VIII.2**).



**IV.2. Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking.**

---



# Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking

Floris Tijmen van Vugt<sup>1,2\*</sup>, Hans-Christian Jabusch<sup>3</sup> and Eckart Altenmüller<sup>1</sup>

<sup>1</sup> Institute of Music Physiology and Musicians' Medicine, University of Music, Drama and Media, Hanover, Germany

<sup>2</sup> Lyon Neuroscience Research Center CNRS-UMR 5292, INSERM U1028, University Lyon-1, Lyon, France

<sup>3</sup> Institute of Musicians' Medicine, Dresden University of Music "Carl Maria von Weber," Dresden, Germany

## Edited by:

Gottfried Schlaug, Beth Israel  
Deaconess Medical Center and  
Harvard Medical School, USA

## Reviewed by:

Gottfried Schlaug, Beth Israel  
Deaconess Medical Center and  
Harvard Medical School, USA  
Shinya Fujii, Beth Israel Deaconess  
Medical Center and Harvard Medical  
School, USA

## \*Correspondence:

Floris Tijmen van Vugt, Institute of  
Music Physiology and Musicians'  
Medicine, University of Music, Drama  
and Media, Emmichplatz 1, Hanover  
30175, Germany.  
e-mail: f.t.vanvugt@gmail.com

We investigated how musical phrasing and motor sequencing interact to yield timing patterns in the conservatory students' playing piano scales. We propose a novel analysis method that compared the measured note onsets to an objectively regular scale fitted to the data. Subsequently, we segment the timing variability into (i) systematic deviations from objective evenness that are perhaps residuals of expressive timing or of perceptual biases and (ii) non-systematic deviations that can be interpreted as motor execution errors, perhaps due to noise in the nervous system. The former, systematic deviations reveal that the two-octave scales are played as a single musical phrase. The latter, trial-to-trial variabilities reveal that pianists' timing was less consistent at the boundaries between the octaves, providing evidence that the octave is represented as a single motor sequence. These effects cannot be explained by low-level properties of the motor task such as the thumb passage and also did not show up in simulated scales with temporal jitter. Intriguingly, this instability in motor production around the octave boundary is mirrored by an impairment in the detection of timing deviations at those positions, suggesting that chunks overlap between perception and action. We conclude that the octave boundary instability in the scale playing motor program provides behavioral evidence that our brain chunks musical sequences into octave units that do not coincide with musical phrases. Our results indicate that trial-to-trial variability is a novel and meaningful indicator of this chunking. The procedure can readily be extended to a variety of tasks to help understand how movements are divided into units and what processing occurs at their boundaries.

**Keywords:** piano scale, variability, motor sequence, chunking, perception

## INTRODUCTION

Playing music means executing particular motor commands to create an auditory stimulus (Jäncke, 2012). Both the motor system and musical materials are highly structured. How can we gain insight into these two kinds of structure: motor sequences and musical phrases? And, more interestingly, how do they interact? Understanding this fundamental problem is crucial to understanding how perception and action are related.

Insight into the organizational structure of motor actions is provided by sequence learning paradigms. Participants learning to type a sequence of numbers divide it into smaller sub-sequences so as to facilitate learning (Koch and Hoffmann, 2000). Keystrokes at the boundaries of these units or "chunks" (Sternberg et al., 1978; Soechting and Terzuolo, 1987; Hikosaka et al., 1995) are robustly found to be played slower than chunk-internal keystrokes (Povel and Collard, 1982; Rosenbaum et al., 1983; Sakai et al., 2003). Probably, this slowing down is a side-effect of the increased cognitive load of having to transition between the sequence chunks, which in themselves can be executed in an automated, feed-forward fashion (Lashley, 1951).

On the other hand, the musical material is thought to contain structural cues in timing deviations that are, intriguingly, reminiscent of the timing effects found in sequence production literature. That is, pianists slow down at the end of musical phrases (Palmer and Krumhansl, 1987; Repp, 1992a; Furuya et al., 2011). Such phrase-final slowing persist even when instructed to play "metronomically" (Repp, 1999b), and the more advanced a pianist the larger such deviations tend to become (Sloboda, 1983). Tempo variations thus appear to be a communicative device to signal structural landmarks of the musical piece. The process is altogether reminiscent of the widespread phrase-final lengthening in natural language (Klatt, 1976; Turk and Shattuck-Hufnagel, 2007).

Do these results imply that the structure of musical phrases is that of the motor commands used to create it? In other words, could musical phrases and motor sequences be two sides of the same coin? This idea conflicts with the *perceptual hypothesis* (Drake, 1993), which explains temporal deviations in music production as resulting from perceptual biases. For example, suppose that the last interval in an isochronous sequence always sounds too short. In other words, our perceptual timing space is warped (Penel and Drake, 1998). As a compensatory strategy, pianists, aiming for

perceptual regularity, rather than “objective” regularity, will gravitate toward lengthening that final interval (Repp, 1999b). In brief, biases in perception likely cause training to favor certain irregular patterns rather than objective evenness.

Then what causes temporal deviations in musical playing: motor chunking or perceptual deviations? We propose musical scale playing as our paradigm to investigate this question. Scales are the quintessential musical practice materials: the first thing Mozart’s *Zauberflöte* character Tamino plays on his flute is a scale. Indeed, practising scales is one of the chores that every classical pianist is engaged in for many hours during their professional career. As a result, we can expect that their motor structure has become sufficiently stable. In order to disentangle motor sequence structure and musical phrasing, we present a novel analysis of scale timing. This method segments the unevenness of playing into:

- (i) systematic deviations from objective evenness that are perhaps *residuals of expressive* timing (Repp, 1999a) or resulting from perceptual biases (Penel and Drake, 1998), which we refer to as irregularity, and
- (ii) non-systematic deviations that can be interpreted as motor execution errors, perhaps due to noise in the nervous system (Harris and Wolpert, 1998; Faisal et al., 2008), which we will call *instability*.

The analysis is described in more detail below.

In this study we will first of all present a validation of our irregularity-instability analysis by showing that it allows one to reconstruct the previously used unevenness measure (Experiment I). We hypothesize that instability of the various notes in a scale will indicate how the motor program is chunked, whereas irregularity reveals its musical structure (Experiment I and II). Finally, in order to gain insight into the link between perception and action, and we investigate the relation between auditory perception resolution on playing instability (Experiment III).

Furthermore, our study is the first to investigate various types of scales, thus being able to control for differences in motor program (i.e., the fingering) and musical content (for example, major vs. minor scales). In addition to the C-major scale, we will include two other scales as controls. The first is the A-minor scale, which is of interest to us since it is played with a fingering identical to C-major. That means, in terms of low-level motor execution it is exactly the same as C-major (except of course for its being played three tones lower than C-major) whereas in terms of the musical content and the tension-resolution profile it is very different (Krumhansl and Kessler, 1982). As a counterpart to this control we took the F#-major scale, which is very different as regards the finger movements, but the relative tone distances between the notes are the same as C-major.

## EXPERIMENT I

### MATERIALS AND METHODS

#### *Experimental set-up*

Thirty-four right-handed pianists were recruited from the student pool at the Hanover University of Music. Participants (17 female) were on average 24.72 (SD 4.47) years old. They started piano training at 6.4 (SD 2.2) years of age, and accumulated an average

of 14140 (SD 8894) practice hours at their instrument. None of these participants reported any neurological disorder or problems related to performing, such as chronic pain. Participants played a Kawai MP9000 stage piano connected to a Pioneer A109 amplifier. The MIDI data was captured through an M-Audio MIDI to USB converter and fed in to a Linux-PC running a custom developed C program that captured the MIDI events. The participants were invited to first play a few minutes to get used to the set-up and warm up. Then they played the scale exercises, which are explained in detail below. The exercises were presented in note score format with indicated (standard) fingering. The pianists were asked to play as regularly as possible at a comfortable mezzo-forte loudness and in legato style. The entire procedure took about half an hour, and the pianists received a nominal financial compensation (10 Euro). The experiment was performed in accordance with the Declaration of Helsinki.

#### **Design**

Participants played two-octave piano scales accompanied by a metronome at 120 BPM. They played four notes within a metronome beat, i.e., eight keystrokes per second. They played blocks of approximately 30 alternating ascending and descending scales with a 9-note rest in between. The scales were played in the following blocks, separated by small breaks: (i) C-major with the right hand, (ii) C-major with the left-hand, (iii) A-minor with the right hand, (iv) F#-major with the right hand, (v) C-major with both hands. The left-hand and both-hand conditions were included as part of our scale playing battery, but will not be reported on in this paper. The C-major and A-minor scales were played with their conventional fingering (123123412312345, where the numbers indicate the fingers from the thumb, 1, to the little finger, 5, and the F#-major with 234123123412312). Following musicological convention, we will refer to the notes by their rank in the scale, in ascending order:  $\hat{1}, \hat{2}, \hat{3}, \hat{4}, \hat{5}, \hat{6}, \hat{7}, \hat{1}', \hat{2}', \hat{3}', \hat{4}', \hat{5}', \hat{6}', \hat{7}', \hat{1}''$ . The C-major scales started at middle c at 262 Hz (c4). The f-minor scale started at 220 Hz (a3) and the F#-major at 370 Hz (F#4).

#### **Establishing scale playing unevenness**

Previous scale playing studies computed the SD of the intervals between the subsequent onsets of the keystrokes as a measure of playing unevenness (Seashore, 1938; Wagner, 1971; MacKenzie and Van Eerd, 1990). More recently, this unevenness metric has been shown to be an indicator for pianistic expertise (Jabusch et al., 2009) as well as for sensorimotor coordination deficits in pianists (Jabusch et al., 2004).

However, a shortcoming of this metric is that it cannot be applied to investigate single-note timing deviations relative to an established temporal reference. For example, suppose one note in the scale is played too late, which is referred to as an “event onset shift” (Repp, 2005). As a result, two intervals are influenced: the interval before this note would come out too long, but the interval following the note would be too short. In brief, a single deviation shows up in two places in an interval-based deviation trace, making it impossible to disentangle the timing of individual notes. Our analysis remedies this problem as described in detail below.

### Scale analysis

First we will describe our analysis of a single played scale. Suppose we have isolated the keystrokes and onsets of one correctly played scale. We then convert the note values to their rank in the scale (so for a C-major scale c would have rank 0, d has rank 1, e has rank 2, etc., up to c' with rank 14) and perform a least-square straight line fit to this set of pairs of rank and timing. This allows us to compute for each note the expected time according to this fit and then the deviation of the timing of the actually measured onset (in ms).

Now we turn to our procedure to analyze the entire MIDI recording for a single participant. First we identified correctly played ascending and descending scales. We then performed the analysis described above for each scale separately and group the obtained temporal deviation values by playing direction (ascending or descending) and by note, yielding a set of 30 such deviations, one for each repetition.

## RESULTS

### Irregularity and instability example

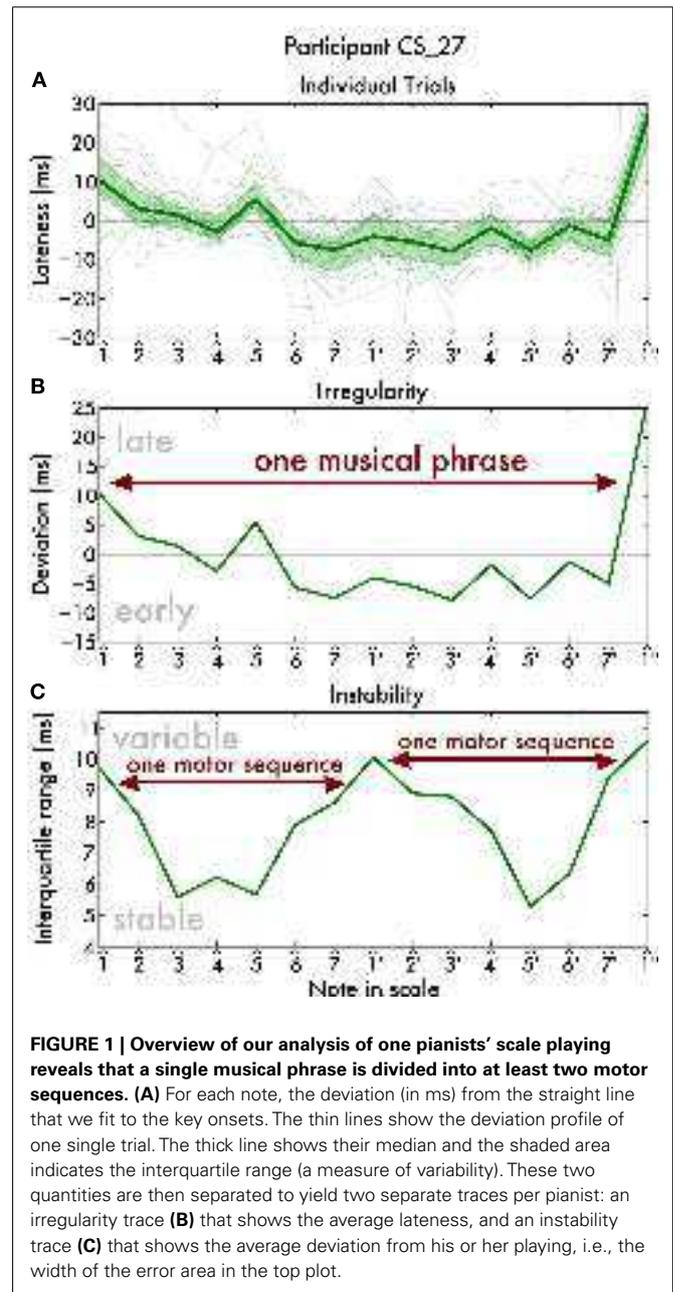
As an illustration of how irregularity and instability were computed, we will present the data of a single participant playing a two-octave C-major scale. Each line in the **Figure 1A** represents a single trial and shows the deviation of each note from the fitted straight line (in ms). By combining the 30 repetitions, we found that some notes are systematically late, such as for example, note 5 (g) is reliably around 8 ms late. We call such systematic deviation *irregularity*. Independently of this, some notes show high trial-to-trial variability, such as for example 1' (c'), whereas other notes show low variability, i.e., they are played very consistently. To quantify this, we calculate the interquartile range of the note timings and refer to this as *instability*. That is, at each note position, irregularity was quantified as the mean timing deviation, and instability as the interquartile range, across all 30 trials.

Irregularity corresponds to what most previous studies have investigated. We found the irregularity trace is roughly an arc (**Figure 1B**), in agreement with widespread findings in the timing literature (Palmer and Krumhansl, 1987; MacKenzie and Van Eerd, 1990; Friberg et al., 2006). That is, playing is slower in the beginning, then speeds up and finally slows down at the end. We conclude from this that the two-octave scale is played as a single musical phrase. Surprisingly, the instability trace reveals a different picture, showing three distinct peaks that coincide with the octave boundaries (**Figure 1C**), instead of two peaks in the irregularity trace.

In sum, we show that, although the two-octave scale is played as a single musical phrase, it is divided into two motor sequences with higher instability at their transition. Our study is the first that we are aware of to reveal this separation.

### Extraction of correctly played scales

One participant was eliminated because he did not follow the instructions to play in a legato style. Scales that were incorrectly played were rejected (3.9% of the note onset events) as were scales for which the least-square fit had an  $R^2$  less than 0.9 (0.07%), which reflected anomalies in playing. One participant's F#-major scales were rejected altogether because he consistently played b# instead of b. Data from the correctly played scales was used to calculate



**FIGURE 1 | Overview of our analysis of one pianist's scale playing reveals that a single musical phrase is divided into at least two motor sequences. (A)** For each note, the deviation (in ms) from the straight line that we fit to the key onsets. The thin lines show the deviation profile of one single trial. The thick line shows their median and the shaded area indicates the interquartile range (a measure of variability). These two quantities are then separated to yield two separate traces per pianist: an irregularity trace (**B**) that shows the average deviation from his or her playing, i.e., the width of the error area in the top plot.

the irregularity and inconsistency for each pianist, exercise, and note separately. The least-square fit to the scales had an average  $R^2$  of 0.9996. The procedure yielded an average of  $31.2(\pm 3.09)$  correctly played ascending and descending scales per exercise per participant.

### Comparison with standard deviation of inter-onset intervals

First, we compare our irregularity-instability analysis to the existing measure of unevenness. For each participant, scale, hand, and playing direction (ascending, descending), we computed unevenness, irregularity, and instability as follows. The unevenness was calculated by taking the SD of the inter-keystroke-intervals in each scale run and then averaged for all runs in each playing direction.

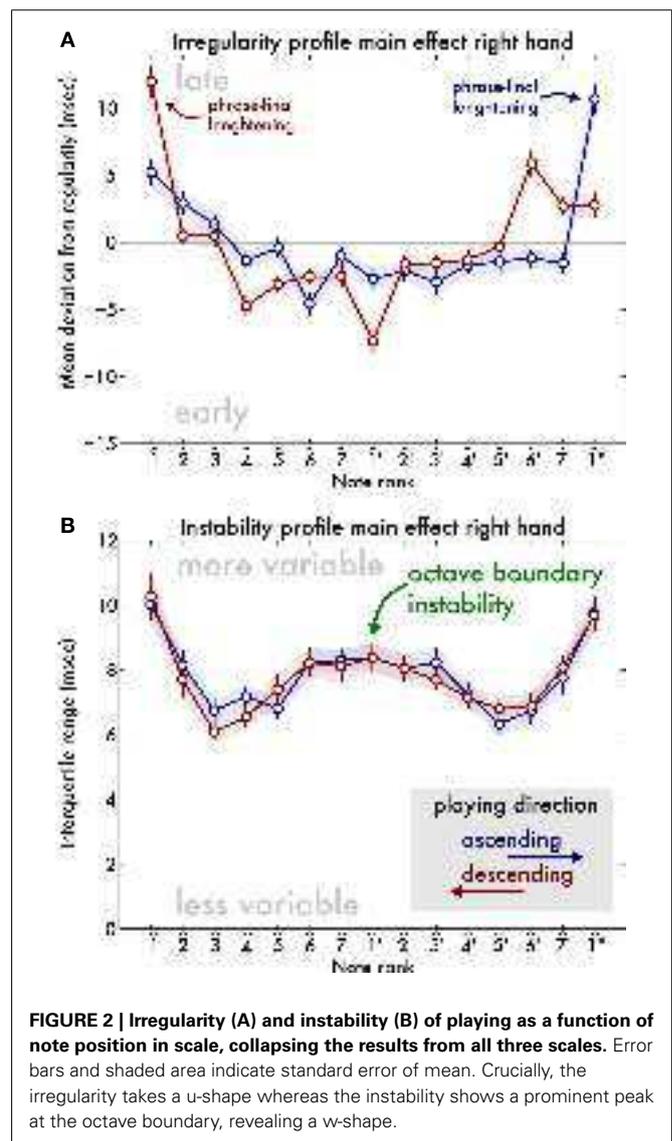
Irregularity was computed as above for each note in the scale and then averaged across runs in each playing direction. Instability was calculated as the interquartile range of the deviations of each note, and then averaged across the notes in each scale and playing direction. That is, for each participant we obtained six scale conditions (the five scale tasks listed in the methods, one of which was played bimanually) times two playing directions (ascending, descending), that is 12 data points, for each of which we had three scalars: the unevenness, irregularity, and instability. We then proceeded with a multiple linear regression to predict the former on the basis of the latter two. Both irregularity and instability resulted as significant factors (both  $t > 16.0$ ,  $p < 0.001$ ). The adjusted  $R^2$  of the obtained model is 0.81. Taking either of the two factors alone yields worse  $R^2$  of 0.69 and 0.61, respectively. The two-factor model is significantly better than the single-factor models with irregularity [ $F(1,32) = 254.56$ ,  $p < 0.001$ ] or instability [ $F(1,32) = 416.64$ ,  $p < 0.001$ ] only. Our model is summarized in the equation  $IOI-SD = 0.93 \cdot \text{irregularity} + 0.64 \cdot \text{instability} + 0.62$ . In brief, using our irregularity-instability analysis, we can reconstruct the unevenness measure with high precision.

### Instability across notes

First, we performed an overall three scales (C-major, A-minor and F#-major)  $\times$  2 directions (ascending, descending)  $\times$  15 notes ANOVA with irregularity as an outcome variable, and the participants as error terms. We found no main effect of scale [ $F(2,62) = 1.46$ ,  $p = 0.24$ ], nor of direction [ $F(1,31) = 0.12$ ,  $p = 0.73$ ]. There was, however, a main effect of note position [ $F(14,434) = 42.50$ ,  $p < 0.001$ ], indicating that note timing varied across the scale. Furthermore, an interaction of scale with note [ $F(28,868) = 6.10$ ,  $p < 0.001$ ] indicated that this timing trace differs across the three scales under investigation. There was no interaction between scale and direction, but there was between direction and note [ $F(14,434) = 12.27$ ,  $p < 0.001$ ] as well as a three-way interaction between scale, note, and direction [ $F(28,868) = 6.03$ ,  $p < 0.001$ ]. **Figure 2A** shows the main effect of note position for the two directions separately, collapsed across the three scales, revealing a u-shaped curve as in our example **Figure 1B**.

Second, we performed the same ANOVA (3 scales  $\times$  2 directions  $\times$  15 notes) but with instability as an outcome measure. We found no main effects of scale or direction, but again a main effect of note [ $F(14,434) = 48.21$ ,  $p < 0.001$ ], indicating that note timing variability was different at different positions in the scale. Of the two-way interactions, only scale with note was significant [ $F(28,868) = 2.32$ ,  $p < 0.001$ ] and the three-way interaction scale, note, and direction was significant [ $F(28,868) = 2.14$ ,  $p < 0.001$ ]. **Figure 2A** shows the main effect of note position on irregularity for the two directions separately but collapsed across the three scales. Instability, however, showed a qualitatively different, w-shaped trace (**Figure 2B**).

We used trend analysis using orthogonal polynomials to investigate the contributions of the various polynomial degrees to the main effect of note on instability. We found no linear or cubic effect but a quadratic (u-shaped) and quartic (w-shaped) effect [ $F(1,31) = 110.24$ ,  $p < 0.001$  and  $F(1,31) = 511.87$ ,  $p < 0.001$ , respectively] as well as significant fifth- and eighth-degree contributions [ $F(1,31) = 4.3$ ,  $p = 0.04$  and  $F(1,31) = 8.56$ ,  $p < 0.001$ ].



**FIGURE 2 | Irregularity (A) and instability (B) of playing as a function of note position in scale, collapsing the results from all three scales.** Error bars and shaded area indicate standard error of mean. Crucially, the irregularity takes a u-shape whereas the instability shows a prominent peak at the octave boundary, revealing a w-shape.

The dominant trend is by far the quartic trend, corresponding to 78.92% of the total sum of squares of the note effect. Indeed, *post hoc* paired *t*-tests on the C-major scale reveal greater instability at the beginning of the two-octave scale [instability at  $\hat{1}$  is greater than at  $\hat{4}$ ,  $t(32) = 9.5$ ,  $p < 0.001$ ] as well as at the end [instability at  $\hat{1}''$  is greater than at  $\hat{5}'$ ,  $t(32) = 7.3$ ,  $p < 0.001$ ]. Surprisingly, a third peak of instability was found at the transition between the two-octaves [instability at  $\hat{1}'$  is greater than at  $\hat{4}$ ,  $t(32) = 6.1$ ,  $p < 0.001$  and also than at  $\hat{5}'$ ,  $t(32) = 6.3$ ,  $p < 0.001$ ]. Instability was the same at the beginning and end of the scale [instability at  $\hat{1}$  and  $\hat{1}''$  are not significantly different,  $t(32) = -0.8$ ,  $p > 0.4$ ].

At this point, one may wonder whether there are systematic timing differences at the octave boundary. That is, does the peak in instability at the middle octave boundary also appear in the irregularity trace? Trend analysis using orthogonal polynomials in the irregularity trace, revealed a strong quadratic trend [ $F(1,31) = 899.24$ ,  $p < 0.001$ ], in line with visual observation of the u-shape of **Figure 2A**, as well as a cubic trend [ $F(1,31) = 27.37$ ,

$p < 0.001$ ]. However, a quartic (w-shaped) trend was found as well [ $F(1,31) = 49.79, p < 0.001$ ], as well as various higher order polynomials. In this case, the dominant trend is the quadratic trend, containing 82.18% of the variance, whereas the quartic trend only amounts to 4.55%. Indeed, using *post hoc* *t*-tests, we find irregularity it is not greater at  $\hat{1}'$  than at  $\hat{4}$  [ $t(32) = -0.06, p > 0.5$ ] or at  $\hat{5}'$  [ $t(32) = -7.1, p = 1$ ]. However, a strong deviation from regularity occurs at the end of the scale, where the deviation at  $\hat{1}''$  (9.3 ms) is greater than at  $\hat{5}'$  (1.1 ms) for ascending scales [ $t(32) = 3.3, p < 0.005$ ]. In other words,  $\hat{1}''$  is played systematically late. Similarly, in descending scales irregularity at  $\hat{1}$  is greater than at  $\hat{4}$  [ $t(32) = 9.4, p < 0.001$ ]. This is a common finding in the timing literature: musicians slow down at the end of musical phrases (e.g., Repp, 1999a).

Now we turn to the A-minor scale (Figure 3B). In both ascending and descending scales we found the octave boundary instability peak [instability at  $\hat{1}'$  is greater than at  $\hat{4}$ , both  $ts(32) > 3.1, ps < 0.001$  and also comparing to  $\hat{5}'$ , both  $ts(32) > 3.1, ps < 0.003$ ] (Figure 3C). We can conclude that the octave boundary instability is also present in the A-minor scale.

In the F#-major scale we found again the instability peak at the octave boundary in the descending scales [ $ts(32) > 3.9, ps < 0.001$ ] but not for the ascending scales [ $t(31) = -0.06, p = 0.73$  and  $t(31) = 2.54, p = 0.1$  for the second octave].

## DISCUSSION

First of all, we have validated our analysis method. We have decomposed the variability from a single variable (unevenness) into two mostly independent and qualitatively different factors (irregularity and instability). This is comparable to the way a vector in the Euclidian plane can be written as a linear combination of two basis vectors.

The participants' trial-to-trial variability profiles (instability) show a clear w-shape pattern across the two scales, with greater instability at the beginning and end of the scale, but surprisingly also in the middle, at the boundary between the octaves. To our knowledge, our study is the first to reveal such subtle but robust differences in timing consistency. Of course, the irregularity and instability curves are related: when the mean deviation is high, the variance typically also increases. This could explain how instability peaks at the beginning and end of the two-octave scale are

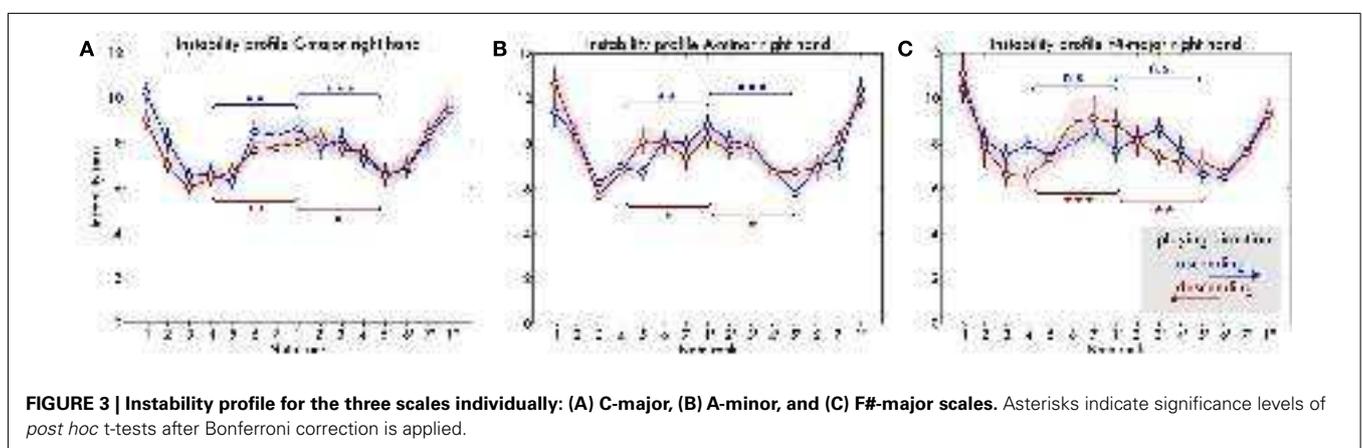
accompanied by irregularity peaks at those locations. This means that our finding of the instability peak at the octave boundary is all the more striking since the u-shaped irregularity curve is at its low-point there.

One may argue these two peaks could alternatively be explained by the mechanical effect of inverting the wrist movement, which switches at those locations between left-to-right and right-to-left movement account, first of all, this movement direction inversion is not abrupt, since ascending and descending scales in our measurement are separated by a 9-note rest (1.125 s). Secondly, such an explanation could not explain why a comparable peak occurs in the middle of the two-octaves.

One other potential explanation for the w-shaped variability pattern is that at the boundary between the two-octaves another event occurs: the thumb passes underneath the fourth finger to be able to play the  $\hat{1}'$  key (in the C-major and A-minor scales), for the inverse playing direction a cross-over maneuver of the fourth finger over the thumb is required. Perhaps these are particularly difficult movements to perform, which would explain the higher variability. For example, a thumb-under movement is accompanied by a substantial horizontal wrist translation that requires preparation (Engel et al., 1997) and a coordinated effort across the different fingers (Furuya et al., 2011). However, we argue these phenomena cannot explain our data sufficiently, because a similar movement occurs in the first octave when also the thumb passes underneath the middle finger (at  $\hat{3}$ ) to play  $\hat{4}$ . And we do not observe increased instability at this note relative to surrounding notes. Furthermore, the very last note,  $\hat{1}''$ , is played without a thumb passage at all and still shows a striking increase in instability. Finally, the F#-major scale contains thumb passages that are much less awkward and in different positions than the C-major scale. Therefore, the w-shape pattern in the descending F#-major scales is not explicable on the basis of thumb passages. We feel these arguments rule out low-level motor explanations of our results.

Another potential explanation for the w-shaped instability would be that this pattern is related to the metronome. However, note that a metronome click occurs every four notes, that is, at  $\hat{1}, \hat{5}, \hat{2}'$  and  $\hat{6}'$  in the C-major scale. These do not coincide with the w-shape instability, making this explanation untenable.

Then the question arises how the increased instability at the octave boundary is to be explained. If the motor system would



conceive of the two-octaves as a single motor program there would be no reason for the playing to become increasingly variable in the middle. Rather, our interpretation is that the octave boundary marks the transition between concatenated motor programs. Under this view, the increased instability is a result of having to load the next sequence into the motor buffer (Lashley, 1951). In other words, the two-octaves are “chunked” into at least two units in the motor system. Again, although we argue that our results indicate segmentation into at least two motor programs, we cannot conclude that the two-octaves make up only two motor programs. It is conceivable that the two-octaves are further subdivided in ways that are not reflected in playing instability, perhaps in a hierarchical fashion (Rosenbaum et al., 1983; Hard et al., 2011).

It is interesting to note that the scales under investigation have revealed the octave boundary effect for both ascending and descending scales (except in the F#-major scales), suggesting that the motor system has chunked these in the same way. Ascending and descending scales are essentially the same movement, but mirrored in time. Therefore, our finding suggests that motor program chunking would be mostly invariant to temporal inversion.

At this point it remains possible, at least in principle, that this instability effect is an artifact of our line fitting procedure or another aspect of our analysis. To control for this, we run the same procedure with simulated data in Experiment II.

## EXPERIMENT II

### MATERIALS AND METHODS

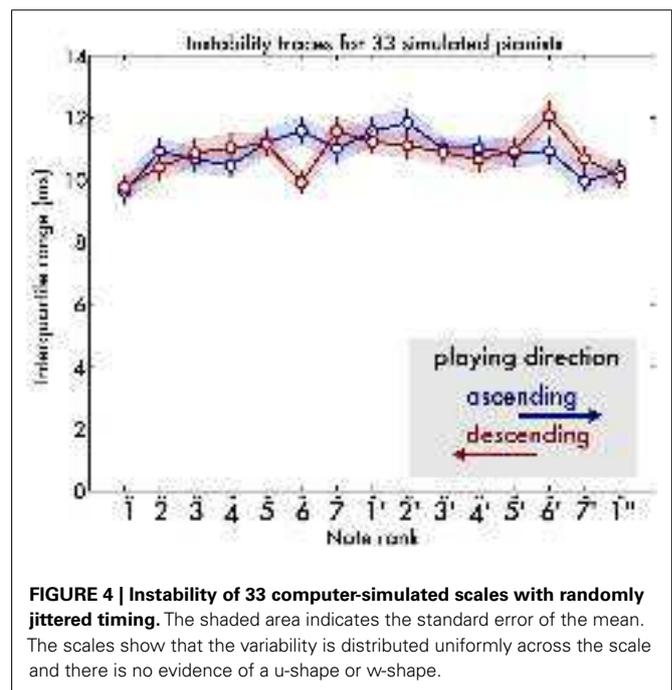
We used a python script to simulate the scale playing of 33 pianists. Each simulated pianist played 30 ascending and descending two-octave C-major scales at 8 notes/s, yielding a total of 900 note onsets that were perfectly regular in time. The timing of each note was then jittered by a time value sampled from a normal distribution with zero mean and a SD of 9 ms. This value was chosen such that the resulting instability profile was on average similar to that found in the real pianists. The same analysis as described for Experiment I was then applied to these data.

### RESULTS

Overall instability levels were comparable to those of the human pianists reported in Experiment I and are shown in **Figure 4**. Although statistics are not commonly computed for simulated data, we reproduce the same analyses as with the human data in order to understand which effects could be due to random variability. We performed a 2 directions  $\times$  15 notes within-participants ANOVA with irregularity as the outcome variability. None of the main effects or interactions were significant (all  $F < 1.5$ ). Performing a similar analysis on instability, we find a main effect of note [ $F(14,448) = 3.26, p < 0.001$ ]. Trend analysis revealed this effect to be mostly due to a quadratic trend [ $F(1,32) = 27.30, p < 0.001$ ], but, crucially, the quartic trend was not significant [ $F(1,32) = 1.81, p = 0.18$ ]. This is further confirmed by our *post hoc* *t*-tests [ $\hat{1}'$  is not greater than  $\hat{4}$ ,  $t(32) = 1.6, p > 0.05$ , similarly  $\hat{1}'$  is not greater than  $\hat{5}'$ ,  $t(32) = 1.0, p > 0.15$ ].

### DISCUSSION

We ruled out the possibility that the octave boundary instability that was seen in the recordings of pianists would be an artifact to our analysis. If so, it would also have occurred in the simulated corpus.



**FIGURE 4 |** Instability of 33 computer-simulated scales with randomly jittered timing. The shaded area indicates the standard error of the mean. The scales show that the variability is distributed uniformly across the scale and there is no evidence of a u-shape or w-shape.

Our novel analysis of scale playing has divided the playing unevenness into largely independent components: systematic deviations from regularity and trial-to-trial instability. *A priori*, one would expect that only the former, systematic deviations could be the result of a compensation mechanism that plays longer temporal intervals that sound shorter and vice versa (Drake, 1993). The reason is that such perceptual biases should be the same between trials and therefore give rise to consistent temporal deviations. Therefore, the instability peaks cannot be the result of perceptual warping. In other words, we expect the detection profile to correlate with the produced irregularities and not the instabilities.

In order to test this hypothesis we performed a temporary delay detection experiment in which one note was delayed at various positions in the scale, and participants were asked to detect this. Since Experiment I revealed the octave boundary to be present in the three scales (although to a lesser extent in F#-major), we decided to restrict our current investigation to the C-major scale only.

## EXPERIMENT III

### MATERIALS AND METHODS

Nineteen music students of various instruments were recruited from the Hanover University of Music. We used a python script to generate two-octave C-major scales (from  $\hat{1}$  up to and including  $\hat{1}''$ ), ascending only. The tones were played perfectly regularly at 8 notes/s except for one of five notes ( $\hat{1}$ ,  $\hat{4}$ ,  $\hat{1}'$ ,  $\hat{4}'$ , or  $\hat{1}''$ ) that was delayed by 40 ms. This procedure is similar to previous perceptual studies (Ehrlé and Samson, 2005). Five additional two-octave C-major scales with no deviation were inserted to yield a total of 10 stimuli and the entire set was presented twice in random order. The scales were preceded by two short high-pitch piano tones with a 500 ms interval to establish the temporal reference.

The python script generated MIDI files (using the MXM Python MIDI package), which were then converted offline into wave using Timidity and presented using Audacious. Participants indicated on a paper form for each scale whether they heard a timing deviation. As a training, they first heard two example scales with a (longer) deviation and two scales without and received accuracy feedback.

## RESULTS

Overall, the participants responded 79% (SD 8.2%) correctly, showing the feasibility of the task. For each of the five delay locations we calculated the hit rate (correct answers/number of presentations). It is not possible to calculate a  $d'$ -score since the presentation of the five delay locations were randomly interleaved.

Detection rates were above chance level at all locations in the scale (binomial test all  $p$ s < 0.005) except for, crucially, at the middle octave boundary at  $\hat{1}'$  (binomial  $p > 0.07$ ). Adjacent  $\hat{4}$  and  $\hat{4}'$  were significantly better (Fisher exact  $p = 0.01$  and  $p = 0.003$  respectively). Interestingly, detection rate at the end of the scale ( $\hat{1}''$ ) was worse than chance, meaning that participants were biased to not perceive a perturbation there even if there was (Figure 5).

## DISCUSSION

This experiment reveals that detection accuracy varies by note position in the scale. Overall accuracy was good, showing that the task was feasible. Optimal performance was seen in the middle of either of the octaves (at  $\hat{4}$  and  $\hat{4}'$ ). Detection was at chance level at the beginning of the two-octave scale (at  $\hat{1}$ ) although the participants had a pre-established temporal reference because of the metronome clicks. Detection was also impaired at the end of the two-octave scale, which is a result that ties in well with previous accounts of how perception accuracy is influenced by expectation (Repp, 1999b): listeners expect a slowing down at the end and

therefore when there is a delay in that position the stimulus is not perceived as deviant. This is reflected by the fact that performance was below chance level. If participants responded randomly they would have a detection rate of 50%, but in our case they actually exhibit a bias toward not hearing the deviation, even if there is one.

The crucial case, however, is the transition point between the two-octaves. We observed a decreased auditory sensitivity to delays this point, but pianists' playing shows no systematic deviation (Experiment I). Thus, one must reject our initial hypothesis that the auditory detection profile mirrors the irregularity trace. This is a tantalizing finding that nuances the way perceptual distortions affect action: a loss of perceptual resolution is reflected in a loss of playing stability in the absence of consistent timing deviations (irregularity). Finally, this result again undermines the interpretation that the instability peak at the octave boundary is related to a low-level motor process such as fingering. Such an interpretation would predict that there is no deterioration of auditory perception at the boundary, whereas our experiment shows there is.

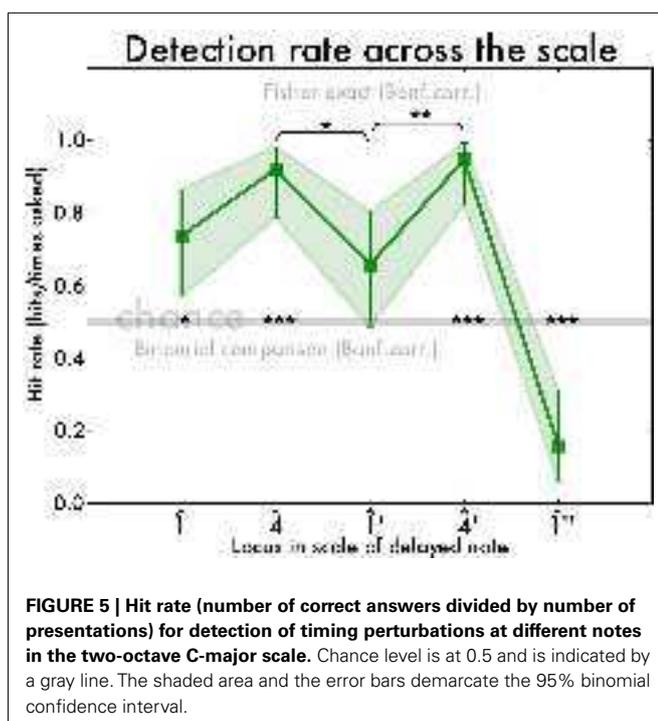
Previous studies have investigated sensitivity to timing changes in regular sequences of events (Hyde and Peretz, 2004; Ehrle and Samson, 2005). Typically, these paradigms employ sequences of five tones of which the fourth is delayed. Listening to deviations in two-octave scales is different since (i) the sequences are longer (15 notes), and (ii) the items in the sequences are not identical, but vary with increasing pitch. Thus, the observed drop of temporal deviation detection at the octave boundary is likely the result of the auditory system's dividing of the temporal sequence in the same, octave-unit chunks as was evident in the motor system. Indeed, previous studies using self-paced viewing of a slide show of actions showed longer looking times at boundaries of actions (Hard et al., 2011), suggesting increased processing demands at the boundaries of perceptual units. In our experiment, however, the notes arrived at fixed intervals beyond the control of participants. Therefore the increased processing load at the chunk boundaries is likely to have interfered with the processing of the items themselves, thus explaining our observed perceptual effect. As such, we predict that the same effect should be observable in the other scales (A-minor and F#-major) as well.

In sum, we have revealed a parallel between the instability trace in pianists' playing and the detection rate of timing perturbations in listeners.

## GENERAL DISCUSSION

### OVERVIEW

Our results indicate that it is possible to meaningfully dissociate irregularity as planned by the motor system and instability of the execution of the motor program. Experiment I revealed that these two factors contribute to the SD of the keystroke intervals as investigated in previous studies. The advantage of our analysis is that we tease apart systematic deviations, which can be rooted in perceptual biases (Penel and Drake, 1998) or residual expressive timing (Repp, 1999a), from mere motor execution errors. Another advantage is that the line fit can still be computed even if a few notes are missing or played twice. As a result, our analysis is robust enough to be applied to pianists who play relatively few correct scales, for example because of a movement disorder such as musician's dystonia (Jabusch et al., 2004) or experimental design (Maidhof et al., 2009).



**FIGURE 5 | Hit rate (number of correct answers divided by number of presentations) for detection of timing perturbations at different notes in the two-octave C-major scale.** Chance level is at 0.5 and is indicated by a gray line. The shaded area and the error bars demarcate the 95% binomial confidence interval.

Note-by-note investigation reveals that instability is greater at the boundaries between the octaves. This is true for two-octave C-major scales and the motorically identical A-minor scales, revealing that it is not related to the C-major musical content but related to motor execution. We interpret these results as revealing that at the octave boundary a transition occurs between subsequent motor program chunks. Previous studies have interpreted the chunking as an aid to learning (Sakai et al., 2004), which is supported by evidence that in the course of learning, the smaller chunks are merged into larger ones (Rosenbaum et al., 1983). In our interpretation, this instability is a side-effect of loading the next motor sequence into the motor buffer (Lashley, 1951). Interestingly, our data stand in contrast to the previous findings of pianists' tapping, showing reduced variability of taps at sequence boundaries (Loehr and Palmer, 2007). However, the latter study used much shorter sequences and calculate the variability of the intervals instead of note-by-note variability. These experimental differences may explain the different and robust findings in the present study.

The octave boundary instability is less strong but still present in the F#-major scales. This can readily be explained by the fact that F#-major scales are much less intensively practised because they are less common in the music literature. For example, one participant did not know it is played with a b rather than a b#. In other words, the F#-major scale may be represented more note-by-note in the motor system because it is played less frequently.

Indeed, we searched the ThemeFinder corpus (<http://www.themefinder.org>, containing 9792 themes at the time of our search). Of these, 936 themes were composed in C-major. Any theme in A-minor revealed 467 matches. F# major themes were only 139 in number. Indeed, these three counts are not independent:  $\chi^2(2) = 624.4$ ,  $p < 0.001$ . In particular, there are between 2.8 and 3.2 (binomial 95% confidence interval) times more C-major themes than A-minor themes. Similarly, there are between 1.2 and 1.4 times (binomial 95% confidence interval) more A-minor than F#-major themes. Of course, some caution is needed in interpreting these corpus search results. They only concern themes and not entire pieces, which may contain modulations. Furthermore, the data is from classical and baroque periods only, and are therefore not necessarily the same as the typical pianists' repertoire. However, it is likely that the distributions of tonalities are at least comparable. Furthermore, this difference between F#-major and C-major scales cannot be explained by washing out of the instability differences due to higher overall instability in the former case, because the mean instability is the same in both cases (Figure 3).

In order to further clarify the processing that occurs at the octave boundary, future investigation could add a weight to the wrist during scale playing, increasing its inertial mass. This means that the preparation for the thumb passage movement would have to be longer (Engel et al., 1997) and likely accompanied by more variability. Therefore, we predict the appearance of two additional peaks in the instability (at  $\hat{4}$  and  $\hat{4}'$ ). However, great caution should be taken in such an endeavor, since the pianist's muscular system is highly sensitive to changes in the playing environment (Sakai, 2002) and can easily result in injury. Alternatively, TMS could be applied during the production of the first octave, which would likely interrupt the output that is currently in the motor buffer.

However, the second chunk (octave) is at that point not yet in the motor buffer and should therefore not be affected.

The picture that emerges from Experiment I is that the octave boundary instability is mainly a low-level motor sequencing phenomenon. But Experiment III reveals an unexpected parallel in perception: that the detection rate is lower at the boundaries of the two-octaves. Assuming that the two phenomena are causally related, which one is the cause and which the effect?

First we consider the possibility that the lack of auditory resolution at the octave boundary causes the playing to be less precise at those points. Similar hypotheses have been advanced that relate musical production to perception. One is that slowing down at the end is musically appropriate and not perceived as deviant (Repp, 1992b), and therefore played that way too. On another interpretation the perceptual space is non-veridical, with certain intervals (such as the last intervals of phrases) sounding shorter and therefore played longer: the perceptual hypothesis (Penel and Drake, 1998). In our case, the problem with both is that they cannot account for the detection impairment at the octave boundary, since we show that there is an increase in playing instability, but no systematic slowing (which would have appeared in the irregularity trace).

The perceptual hypothesis could be amended to include this possibility. Imagine that due to the lack of perceptual resolution at the octave boundary, the interval does not sound systematically shorter, but sometimes shorter and sometimes longer. As a result, the playing would sometimes compensate by playing it longer and sometimes by playing it shorter, yielding increased playing variability but not systematic deviation, in line with our findings. What is not satisfying about this explanation is that it does not account for why perceptual resolution is lower at such locations that do not seem musically meaningful. However, an even more immediate problem is that it would predict the octave boundary instability to be present equally in the F#-major scales, contrary to our findings.

A second explanation is the inverse causality: a lack in playing precision leads to impaired perception. Indeed, participants in Experiment III were musicians and could therefore be heavily influenced by exposure to musical material. A future study could decide this issue by performing the perception experiment on non-musicians. A limitation of such investigation will be that even non-musicians have much been passively exposed to music.

In sum, then, we conclude that the chunks formed in the motor system and in the perceptual system overlap, at least for the materials presently studied. How chunk formation in the two systems is causally related remains yet to be answered.

These questions open the road to further investigation into the relation between music perception and production, which may be more complex than previously accounted for. A future study may use a signal-detection-theoretical framework to tease apart response bias and sensitivity and correlate these to playing irregularity and inconsistency. Our prediction is that the irregularity trace will be mirrored by the detection bias, whereas the instability trace reflects the inverse of the sensitivity.

In sum, our study points to a dissociation between musical phrases and motor programs. Musical phrases have previously been found to be indicated by systematic slowing at the end (i.e., increased irregularity), whereas our finding is that motor sequences are demarcated by increased playing instability. Perhaps

the two reflect the previously discovered dissociation between timing processes and item sequencing (Pfordresher, 2003). This issue could be further clarified by testing whether increased keystroke variability is found at the boundary of learned sequences in a serial reaction time paradigm (Nissen and Bullemer, 1987), as our interpretation would predict.

One limitation in the current study is that although contrary to previous studies we have included scales of different tonalities, still all our material consisted of two-octave scales. A future study could investigate scales over three octaves, although caution would need to be taken to control for the larger distance the arm needs to cover to reach the three octaves.

## REFERENCES

- Drake, C. (1993). Perceptual and performed accents in musical sequences. *Bull. Psychon. Soc.* 31, 107–110.
- Ehrle, N., and Samson, S. (2005). Auditory discrimination of anisochrony: influence of the tempo and musical backgrounds of listeners. *Brain Cogn.* 58, 133–147.
- Engel, K. C., Flanders, M., and Soechting, J. F. (1997). Anticipatory and sequential motor control in piano playing. *Exp. Brain Res.* 113, 189–199.
- Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci.* 9, 292–303.
- Friberg, A., Bresin, R., and Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Adv. Cogn. Psychol.* 2, 145–161.
- Furuya, S., Flanders, M., and Soechting, J. F. (2011). Hand kinematics of piano playing. *J. Neurophysiol.* 106, 2849–2864.
- Hard, B. M., Recchia, G., and Tversky, B. (2011). The shape of action. *J. Exp. Psychol. Gen.* 140, 586–604.
- Harris, C. M., and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature* 394, 780–784.
- Hikosaka, O., Rand, M. K., Miyachi, S., and Miyashita, K. (1995). Learning of sequential movements in the monkey: process of learning and retention of memory. *J. Neurophysiol.* 74, 1652–1661.
- Hyde, K. L., and Peretz, I. (2004). Brains That Are Out of Tune but in Time. *Psychol. Sci.* 15, 356–360.
- Jabusch, H.-C., Alpers, H., Kopiez, R., Vauth, H., and Altenmüller, E. (2009). The influence of practice on the development of motor skills in pianists: a longitudinal study in a selected motor task. *Hum. Mov. Sci.* 28, 74–84.
- Jabusch, H.-C., Vauth, H., and Altenmüller, E. (2004). Quantification of focal dystonia in pianists using scale analysis. *Mov. Disord.* 19, 171–180.
- Jäncke, L. (2012). The dynamic audio-motor system in pianists. *Ann. N. Y. Acad. Sci.* 1252, 246–252.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *J. Acoust. Soc. Am.* 59, 1208.
- Koch, I., and Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction time tasks. *Psychol. Res.* 63, 22–35.
- Krumhansl, C. L., and Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol. Rev.* 89, 334–368.
- Lashley, K. (1951). “The problem of serial order in behavior,” in *Cerebral Mechanisms in Behavior: The Hixon Symposium*, ed. L. A. Jeffress (New York: Wiley), 112–146.
- Loehr, J. D., and Palmer, C. (2007). Cognitive and biomechanical influences in pianists’ finger tapping. *Exp. Brain Res.* 178, 518–528.
- MacKenzie, C. L., and Van Eerd, D. L. (1990). Rhythmic precision in the performance of piano scales: motor psychophysics and motor programming. *Atten. Perform.* 13, 375–408.
- Maidhof, C., Rieger, M., Prinz, W., and Koelsch, S. (2009). Nobody is perfect: erp effects prior to performance errors in musicians indicate fast monitoring processes. *PLoS ONE* 4, e5032. doi:10.1371/journal.pone.0005032
- Nissen, M. J., and Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cogn. Psychol.* 19, 1–32.
- Palmer, C., and Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 116–126.
- Penel, A., and Drake, C. (1998). Sources of timing variations in music performance: a psychological segmentation model. *Psychol. Res.* 61, 12–32.
- Pfordresher, P. Q. (2003). Auditory feedback in music performance: evidence for a dissociation of sequencing and timing. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 949–964.
- Povel, D.-J., and Collard, R. (1982). Structural factors in patterned finger tapping. *Acta Psychol. (Amst.)* 52, 107–123.
- Repp, B. H. (1992a). Diversity and commonality in music performance: an analysis of timing microstructure in Schumann’s “Träumerei.” *J. Acoust. Soc. Am.* 92, 2546–2568.
- Repp, B. H. (1992b). Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Cognition* 44, 241–281.
- Repp, B. H. (1999a). Control of expressive and metronomic timing in pianists. *J. Mot. Behav.* 31, 145–164.
- Repp, B. H. (1999b). Detecting deviations from metronomic timing in music: effects of perceptual structure on the mental timekeeper. *Percept. Psychophys.* 61, 529–548.
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* 12, 969–992.
- Rosenbaum, D. A., Kenny, S. B., and Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 86–102.
- Sakai, K., Hikosaka, O., and Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends Cogn. Sci. (Regul. Ed.)* 8, 547–553.
- Sakai, K., Kitaguchi, K., and Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Exp. Brain Res.* 152, 229–242.
- Sakai, N. (2002). Hand pain attributed to overuse among professional pianists. *Med. Probl. Perform. Art.* 17, 178–180.
- Seashore, C. E. (1938). *Psychology of Music*. New York: Dover.
- Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Q. J. Exp. Psychol. A* 35, 377–396.
- Soechting, J. F., and Terzuolo, C. A. (1987). Organization of arm movements. Motion is segmented. *Neuroscience* 23, 39–51.
- Sternberg, S., Monsell, S., Knoll, R. L., and Wright, C. E. (1978). “The latency and duration of rapid movement sequences: comparisons of speech and typewriting,” in *Information Processing in Motor Control and Learning*, ed. G. E. Stelmach (New York: Academic Press), 117–152.
- Turk, A. E., and Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *J. Phon.* 35, 445–472.
- Wagner, C. (1971). “The influence of the tempo of playing on the rhythmic structure studied at pianist’s playing scales,” in *Medicine and Sport Biomechanics II*, eds J. Vredenburg and J. Wartenweiler (Basel: Karger), 129–132.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 April 2012; accepted: 24 October 2012; published online: 16 November 2012.

Citation: van Vugt FT, Jabusch H-C and Altenmüller E (2012) Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Front. Psychology* 3:495. doi: 10.3389/fpsyg.2012.00495

This article was submitted to *Frontiers in Auditory Cognitive Neuroscience*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 van Vugt, Jabusch and Altenmüller. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

**IV.3. Individuality that is unheard of: systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint.**

---



# Individuality that is unheard of: systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint

Floris Tijmen Van Vugt<sup>1,2\*</sup>, Hans-Christian Jabusch<sup>3</sup> and Eckart Altenmüller<sup>1</sup>

<sup>1</sup> Institute of Music Physiology and Musicians' Medicine, University of Music, Drama, and Media, Hanover, Germany

<sup>2</sup> Lyon Neuroscience Research Center, CNRS-UMR 5292, INSERM U1028, University Lyon-1, Lyon, France

<sup>3</sup> Institute of Musicians' Medicine, University of Music "Carl Maria von Weber," Dresden, Germany

## Edited by:

Bruno Gingras, University of Vienna, Austria

## Reviewed by:

Ian E. Holliday, Aston University, UK  
Sarah Creel, University of California San Diego, USA

## \*Correspondence:

Floris Tijmen Van Vugt, Institute of Music Physiology and Musicians' Medicine, University of Music, Drama, and Media, Emmichplatz 1, 30175 Hanover, Germany.  
e-mail: f.t.vanvugt@gmail.com

Whatever we do, we do it in our own way, and we recognize master artists by small samples of their work. This study investigates individuality of temporal deviations in musical scales in pianists in the absence of deliberate expressive intention. Note-by-note timing deviations away from regularity form a remarkably consistent "pianistic fingerprint." First, eight professional pianists played C-major scales in two sessions, separated by 15 min. Euclidian distances between deviation traces originating from different pianists were reliably larger than traces originating from the same pianist. As a result, a simple classifier that matched deviation traces by minimizing their distance was able to recognize each pianist with 100% accuracy. Furthermore, within each pianist, fingerprints produced by the same movements were more similar than fingerprints resulting in the same scale sound. This allowed us to conclude that the fingerprints are mostly neuromuscular rather than intentional or expressive in nature. However, human listeners were not able to distinguish the temporal fingerprints by ear. Next, 18 pianists played C-major scales on a normal or muted piano. Recognition rates ranged from 83 to 100%, further supporting the view that auditory feedback is not implicated in the creation of the temporal signature. Finally, 20 pianists were recognized 20 months later at above chance level, showing signature effects to be long lasting. Our results indicate that even non-expressive playing of scales reveals consistent, partially effector-unspecific, but inaudible inter-individual differences. We suggest that machine learning studies into individuality in performance will need to take into account unintentional but consistent variability below the perceptual threshold.

**Keywords:** piano scale, individuality, expertise, music, recognition

## INTRODUCTION

Our actions are highly individual and we can tell people apart by how they move (Flach et al., 2004; Loula et al., 2005; Prasad and Shiffrar, 2009; Sevdalis and Keller, 2011). People may recognize those close to them by the way they sneeze or walk the stairs. Even when trying to achieve the same aim, the actions that are selected toward this aim and the way in which they are executed vary considerably between individuals. The human observer seems to rely on action simulation to recognize individuals by their movements, since recognition is generally stronger when distinguishing one's own performance from that of others (Jeannerod, 2003).

A first question is how movements from different individuals vary *physically*. Why are certain parameters of our actions remarkably stable between multiple iterations by the same person, and yet strikingly different between individuals? A second question is to what extent movements vary *perceptually*. For example, some movements may differ so subtly that the individual features are not distinguishable to a human observer under normal conditions.

Music is a suitable paradigm to study individuality since actions are directed toward a clearly defined auditory goal: when we play

music, the aim is to make a certain sound. Furthermore, differences between performers are sometimes so salient that listeners will often refuse to listen to a musical piece that is a mere "cover" of the original. Music played by different individuals varies *physically*. For example, machine ensemble learning approaches are able to tell musical performers apart based on structural features such as timing and loudness differences (Stamatatos and Widmer, 2005) or kinematics (Dalla Bella and Palmer, 2011). The individuality is also *perceptual*. Indeed, non-musicians and musicians alike were able to recognize performances reliably (Gingras et al., 2011). Again, action simulation in the form of musical imagery appears to play a role in the recognition process. For example, piano players turn out to be capable of recognizing their own playing from a few months previously, even if the sound was switched off at the time of the recording (Repp and Knoblich, 2004).

In music performance recognition the differences in sound that different players produce are often understood as a result of their artistic individuality. However, there is no reason to assume that the individuality in the way we walk serves any particular purpose. Indeed, even task-irrelevant sounds matching a golf swing are recognized significantly better than chance (Murgia et al., 2012).

On the other hand, individuality in music performance is tacitly assumed to define a performer's unique artistic identity. But we have to date no empirical validation of the extent to which individuality in music performance is deliberate. The study coming closest to answering this question requested pianists to play mechanically, and found that recognition was somewhat impaired for these inexpressive recordings (Gingras et al., 2011). However, even metronomic playing has been shown to contain the same timing patterns as expressive playing, but to a lesser extent (Repp, 1999a). To avoid this problem, we instead investigated the playing of musical scales (Wagner, 1971; MacKenzie and Van Eerd, 1990). When participants are instructed to play a scale as regularly as possible and in a legato style, there is a clear auditory target of perceptual evenness and it is understood that the task at hand is not to play scales in one's own particular way. In other words, isolated scales are not thought of as expressive musical materials. There is some objective standard and trying to meet it is a merely technical task.

Yet, it is found that musical scales show systematic temporal deviations (MacKenzie and Van Eerd, 1990; van Vugt et al., 2012). These deviations are thought of as the result of perceptual distortions (Drake, 1993), residual expressive timing (Repp, 1999a), or of some note transitions involving more difficult movements (Engel et al., 1997).

Our question is whether these temporal deviations are individual in the same way that expressive performance is. We restrict our attention to timing of note onsets, discarding information such as differences in loudness and note duration. In Experiment I, we first established timing deviations of individual notes (van Vugt et al., 2012). The resulting timing profile is then used to recognize pianists across two sessions, separated by 15 min. In this way, we aim to establish individuality that is physically present in the timing of musical scales. In Experiment II, we then proceed to assess whether the timing differences can be perceived by musically trained observers. In Experiment III we investigate the role of auditory feedback in the formation of these timing profiles. Finally, in order to investigate to what extent these timing deviation profiles are stable, we follow a group of pianists over 27 months in Experiment IV.

## EXPERIMENT I

### MATERIALS AND METHODS

The data reported here were collected as part of a validation procedure for a scale unevenness quantification method published elsewhere (Jabusch et al., 2004). Eight pianists (six female) were recruited from the student/teacher pool at the Hanover University of Music and were 24.3 (SD 2.4) years old. All but one were right-handed ( $M = 57.2$ ,  $SD = 66\%$  right-handed according to the Edinburgh handedness inventory). None of the participants reported any neurological condition. Participants played on a MP 9000 MIDI keyboard (Kawai, Krefeld, Germany). The keyboard's digital music interface (MIDI out) signal was captured on a PC using a commercially available sequencer software (MusiCator Win, version 2.12; Music Interactive Technology, Bergen, Norway).

Participants were requested to play two-octave C-major scales beginning with the C (131 Hz) one octave below the middle C and

ending with the C (523 Hz) one octave higher than the middle C. Ascending and descending scales were interleaved. The instruction to the participants was to play as evenly as possible, without expression, and in a legato style at mezzo-forte loudness. A metronome gave a beat at 120 BPM and the instructions were to play at four notes per metronome beat, resulting in eight notes per second. Participants performed 10–15 scales with the right hand and with the left hand (*first measurement*). After a 15 min break, the procedure was repeated (*follow-up*).

### ANALYSIS OF SCALE TIMING

First, we isolated correctly performed scale runs, discarding those containing errors or surplus notes. We then converted the note values to their rank in the C-major scale (i.e., C has rank 0, D has rank 1, E has rank 2, etc., up to  $C''$  with rank 14) and performed a least-square straight line fit to this set of pairs of rank and timing. This allowed us to compute for each note the expected onset time (according to this fit) and then the deviation of the timing of the actually measured onset (in ms) (van Vugt et al., 2012). We performed this fit for all scale runs and then pooled the results by hand (left or right), playing direction (inward or outward) and note, calculating the mean lateness (in ms) for that condition. The result was a  $2$  (hands)  $\times 2$  (directions)  $\times 15$  (notes) matrix of timing deviations, which we will refer to as our irregularity trace. As an illustration, **Figure 1A** shows the irregularity trace for right hand ascending scales in one pianist in the two measurement sessions, and **Figure 1B** for two different pianists. It is clear that the irregularity traces originating from the same pianist (**Figure 1A**) are strikingly similar, whereas those originating from different pianists (**Figure 1B**) are qualitatively different. This is the observation that our analysis (described below) aims to capture.

Additionally, we calculated the unevenness of the scale in accordance with a previously established protocol (Jabusch et al., 2004) as follows. For each correct scale run, the intervals between the consecutive note onsets were calculated and then we took the standard deviation of these. For each hand, direction, and recording (first or follow-up) we took the median of the standard deviations of the scale runs (in ms). The higher this unevenness score, the more temporally irregular the scales.

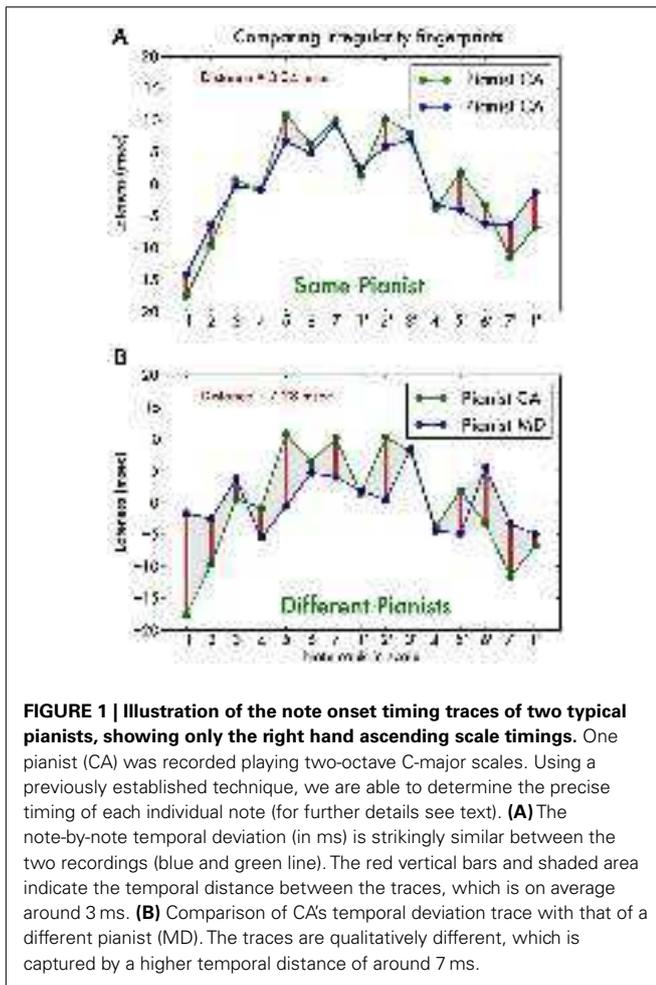
In ANOVAs we report  $\eta_G^2$  as the generalized effect size (Bakeman, 2005). Following musicological notational convention, we will refer to the notes in the scale as 1, 2, 3, 4, 5, 6, 7, 1', 2', 3', 4', 5', 6', 7', 1'', in ascending order.

## RESULTS

### Preliminaries

First, we isolated the correctly played scales, yielding an average total of 11.7 (SD 0.97) scales per person and condition. As a control analysis, we used the number of scales as an outcome measurement in an ANOVA that revealed no significant difference according to hand [ $F(1, 7) = 3.43$ ,  $p = 0.11$ ], direction [ $F(1, 7) \approx 0.00$ ,  $p \approx 1.00$ ], recording session [ $F(1, 7) = 1.19$ ,  $p = 0.74$ ] nor any interaction effect [all  $F(1, 7) < 0.11$ ]. We can conclude that there is no selection bias due to the discarding of scales.

Now we turn to the unevenness measure (the standard deviation of the inter-keystroke-intervals). ANOVA yielded a significant main effect of hand [ $F(1, 7) = 5.73$ ,  $p < 0.05$ ,  $\eta_G^2 = 0.04$ ], showing



that left hand scales were played more unevenly (mean unevenness 9.19 ms, SD 1.67) than right hand scales (mean unevenness 8.44 ms, SD 1.81). This replicates a previous finding (Kopiez et al., 2011). There was no main effect of playing direction [ $F(1, 7) = 0.01, p = 0.92$ ] nor of recording session [ $F(1, 7) = 1.00, p = 0.35$ ] but there was a two-way interaction between direction and recording [ $F(1, 7) = 7.00, p = 0.03, \eta_G^2 = 0.02$ ], showing that although outward scales were played equally evenly across the sessions, inward scales were more even in the follow-up session (unevenness 8.43 ms, SD 1.86) than in the first session (unevenness 9.13 ms, SD 2.33), perhaps revealing a habituation effect.

### Recognizing individual pianists

A salient feature of the temporal traces is that they are highly individual: traces from the same individual but different sessions vary little, whereas traces from different pianists vary much more (Figure 1). To quantify this observation, we define the temporal distance as the Euclidian distance between any pair of vectors representing the irregularity traces. That is, we calculated the sum of squares of the item-by-item distances. Then we divided this by the number of notes in the traces (15 notes for a two-octave scale). Finally, we took the square root to yield a distance value in ms. First we calculate these distances for each of the two hands,

two directions separately. We find that irregularity traces originating from the same pianist have a distance of 3.42 ms (SD 0.89), whereas those originating from different pianists have a distance of 7.24 ms (SD 0.54) (Figure 4). ANOVA with distance as dependent variable shows a significant main effect of self vs. other [ $F(1, 7) = 108.18, p < 0.001, \eta_G^2 = 0.79$ ] but no effect of hand [ $F(1, 7) = 0.55, p = 0.48$ ] nor playing direction [ $F(1, 7) = 0.30, p = 0.60$ ] nor any interaction effect [all  $F(1, 7) < 1.1$ ].

As a result, we designed the simplest possible classification algorithm as follows. Our algorithm is given a database of the irregularity traces for the first measurements of each of the eight pianists. Then it is presented each of the follow-up irregularity traces, without the player label, and its task is to match each pianist to one of the traces in its database. Our algorithm simply chooses the irregularity trace that matches most closely.

This procedure is performed separately for the four sets of average irregularity traces from the two hands and two playing directions. Classification was flawless (100%) for all the right hand scales (inward and outward), as well as the left hand outward scales. In the left hand inward scales, six pianists are classified correctly and two incorrectly. Chance is at 0.125 recognition rate, meaning that in all cases classification is significantly better than chance [binomial  $p < 0.001$ , 95% confidence interval = (0.35, 0.97) for the left hand inward scales and (0.63, 1.0) for the other cases]. When instead of the complete irregularity trace (15 data points per two-octave scale) we used only the unevenness (one data point per two-octave scale) classification rate dropped to between 0.25 and 0.5, which exceeded chance performance only for the right hand inward scales [binomial  $p = 0.01$ , 95% confidence interval = (0.16, 0.84)].

The Euclidian distance is not necessarily the only or best way to quantify the (dis)similarity between irregularity traces. To illustrate this, we perform the same analysis, but this time we compute the correlation (Pearson  $r$ ) between pairs of irregularity traces. ANOVA on the Fisher  $r$ -to- $z$  transformed correlation coefficients shows a main effect of self vs. other [ $F(1, 7) = 63.92, p < 0.001, \eta_G^2 = 0.74$ ], showing that correlations between irregularity traces from the same pianists are higher [ $z(r) = 1.39, SD 0.42$ ] than irregularity traces from different pianists [ $z(r) = 0.40, SD 0.21$ ]. There is no effect of hand except for a trend [ $F(1, 7) = 5.40, p = 0.05, \eta_G^2 = 0.03$ ], nor a main effect of direction [ $F(1, 7) = 2.76, p = 0.14$ ]. Of the interaction effects only that between hand and direction [ $F(1, 7) = 11.50, p = 0.01, \eta_G^2 = 0.10$ ] is significant [all other  $F(1, 7) < 1.05$ ], revealing that whereas left hand traces correlate equally in both playing directions, right hand inward scales correlate higher than outward scales.

We re-ran our recognition algorithm with the only difference that this time, given an irregularity trace to recognize, it chose the irregularity trace that showed the greatest correlation. Recognition rates are identical to those for Euclidian distance: flawless in all but the case of left hand inward scales with six out of two correctly classified (hence still exceeding chance performance).

### Comparing irregularity traces of the same pianist

So far, we have only compared the irregularity traces produced by the same hand and in the same playing direction but by different pianists. How do the traces produced by the same pianist but by

different hands and different directions compare? We argue that these comparisons may provide crucial insight into what causes the timing deviations (Figure 2A). Our reasoning was as follows. If the temporal deviations result from remnants of expressive timing (Repp, 1999a), then we expect irregularity traces that sound similar to be more similar. That is, we expect the left hand inward and right hand outward traces to be closest together (since they have the same auditory result, modulo octave differences), and similarly the right hand outward and left hand inward scales to be close. If, on the other hand, the temporal deviation traces are mostly determined by biomechanical or neuromuscular factors, then we expect traces generated by the same movements to be closer together than those generated by different movements (Figure 2B). More specifically, the pairs of inward and pairs of outward scales are expected to be closer together than pairs with an inward and outward scale.

Furthermore, note that in all these comparisons we have aligned the irregularity traces in time (in the order in which they are played) and not in space (the order in which they appear on the keyboard). That means, when we compare left hand inward and right hand outward scales, they are the same movement in time, but mirrored in space.

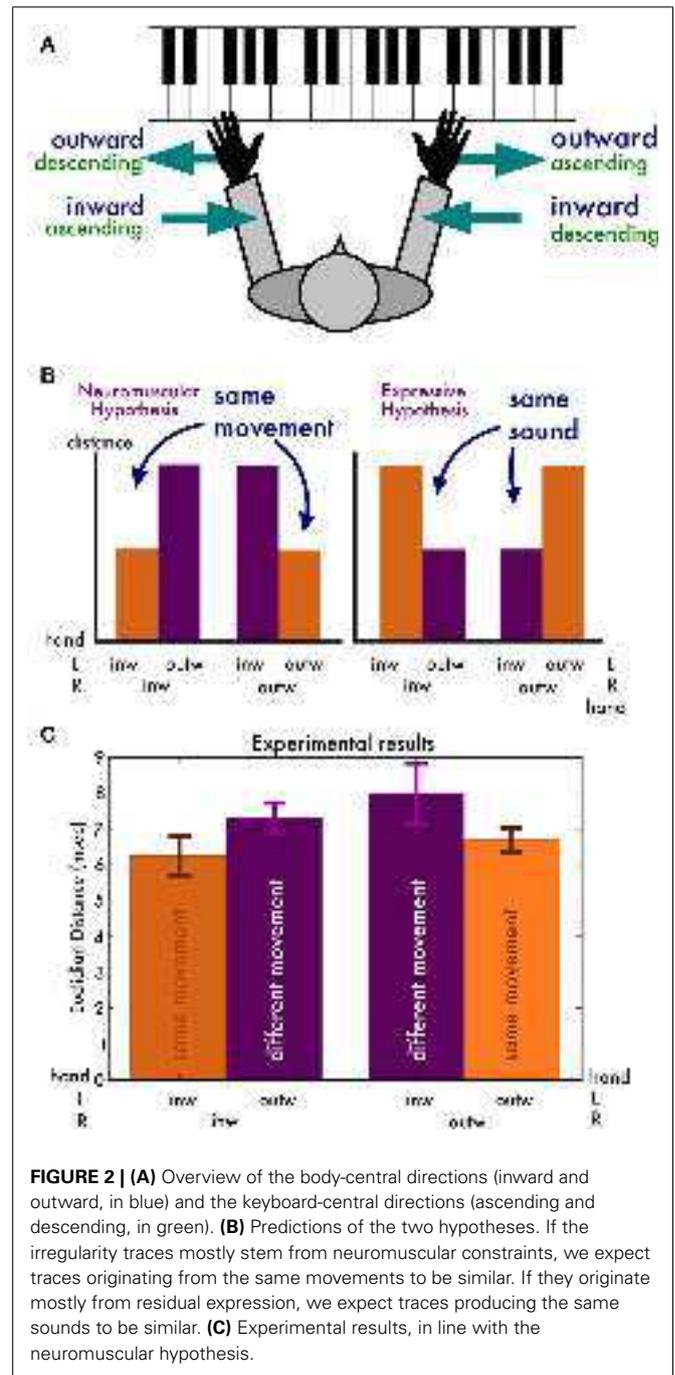
An ANOVA with distance as dependent measure revealed a main effect of movement [ $F(1, 7) = 7.63, p = 0.03, \eta_G^2 = 0.10$ ], reflecting that distances between irregularity traces produced by the same movement are shorter (6.47 ms, SD 0.85) than those produced by different movements (7.64 ms, SD 1.58) (Figure 2C). That is, the results are in line with the hypothesis that the temporal deviations are mostly neuromuscular in nature. No other factor has a main effect [all  $F(1, 7) < 1.6$ ] and there were no interactions [all  $F(1, 7) < 2.0$ ].

**Effector-specificity of the individuality**

To what extent is the individuality in the traces specific to the effector (i.e., hand)? To answer this question, we repeated the analysis above, but comparing the distances across hands within and between pianists. That is, we computed the distance between left and right hand irregularity traces for the same movement direction (inward or outward) and for either the same pianist or different pianists. We found a main effect of same vs. different pianist [ $F(1, 7) = 28.35, p = 0.001, \eta_G^2 = 0.01$ ], revealing that cross-hand distances are smaller between traces from the same pianist ( $M = 6.41, SD = 0.87$  ms) than traces from different pianists ( $M = 7.47, SD = 0.42$  ms). There were no main effects of hand, direction or recording, nor any interaction effects [all  $F(1, 7) < 2.74, p > 0.14$ ].

**DISCUSSION**

Let us pause an instant to take stock. We have shown that pianists do not play scales perfectly regularly. Rather, consistent temporal deviations are present. For the first time we show that these deviations are not mere noise, since they are reliably reproduced across two recording sessions. Furthermore, differences between individuals are so pronounced that a surprisingly simple recognition algorithm is able to recognize pianists nearly flawlessly using the average timing profile of a dozen runs of two-octave scales. The algorithm works equally well when it matches irregularity traces by minimizing distance or by maximizing correlation.



**FIGURE 2 | (A)** Overview of the body-central directions (inward and outward, in blue) and the keyboard-central directions (ascending and descending, in green). **(B)** Predictions of the two hypotheses. If the irregularity traces mostly stem from neuromuscular constraints, we expect traces originating from the same movements to be similar. If they originate mostly from residual expression, we expect traces producing the same sounds to be similar. **(C)** Experimental results, in line with the neuromuscular hypothesis.

An important observation is that the pianists' temporal irregularities are *qualitatively* different. If the irregularity profiles had been qualitatively the same, that is, the same vector simply multiplied by a coefficient, then recognition on the overall unevenness would perform as well as recognition using the entire irregularity trace. But we find the contrary: recognition using a simple overall unevenness metric (the median of the inter-keystroke-intervals) was barely above chance. We can conclude that it is the qualitative differences in the scale timing that enable us to tell the different pianists apart. Hence we can speak of a *pianistic fingerprint*.

What determines this temporal fingerprint? We showed that temporal irregularity traces generated by the same movement are more similar than those generating the same sound. As a consequence, the contribution of biomechanical constraints to these timing profiles must be stronger than expressive or perceptual influences. Furthermore, we found that the individuality in the traces is to some extent effector-independent: the two hands of the same pianist are less different than hands of different pianists. This suggests that the individuality is represented in cortical areas accessible to both effectors (Rijntjes et al., 1999).

In sum, temporal differences are physically present in the produced timing in musical scales. At this point, it remains unclear whether this individuality is also perceptually present: are human observers able to identify performers in the same way our algorithm could?

**EXPERIMENT II  
MATERIALS AND METHODS**

Our perceptual experiment comprised two parts. In the first part (*recognition*), listeners (see details below) were presented with pairs of fingerprint recordings and asked to judge whether they originated from the same or different pianists. Essentially, participants were given the same task that our algorithm in Experiment I performed. In the second part (*irregularity threshold*), we investigated whether participants were able to pick up the temporal irregularities at all by establishing their psychophysical threshold for temporal irregularity. That is, participants were presented a single scale and had to judge whether it was regular (isochronous) or irregular.

**Recognition test**

We took the irregularity traces for the right hand ascending scales for three pianists (CA, ES, and TY) from the *first* and *follow-up* measurements in Experiment I. For each, we furthermore choose one alternative pianist from the *follow-up* measurements (MD, IM, and VH, respectively). Each stimulus consists of a pair of scales played one after the other. These six scale pairs are listed in **Table 1**. Participants responded by pressing a button whether they felt the two scales were played by the same pianist or different pianists.

The two scales in a pair were played preceded by two high-pitched notes (MIDI note 96), providing a tempo reference

at 120 BPM. The scales were then played with four notes per metronome click, that is, at eight notes per second. The second scale always started 3.5 s after the first. All notes had a duration of 137.5 ms to generate legato style and a standardized loudness level. That is, we removed all loudness cues as well as articulation. Furthermore, each scale pair came in two versions: a *veridical* rendition, and a *magnified* rendition where all timing deviations were increased by a factor 5 (for a similar strategy in the context of a recognition experiment, see Hill and Pollick, 2000). In other words, we multiplied the irregularity vector by a scalar, making the differences more salient. The six stimuli (**Table 1**) were rendered twice (veridical and magnified), and presented in the two possible orderings, yielding 24 stimuli. Each of these were presented six times, yielding a total of 144 stimuli. The order was randomized for each participant and divided into 4 blocks of 36 trials.

For data analysis, we used the R Package for Statistical Computing and the signal detection scripts developed by Prof. Abby Kaplan (<http://home.utah.edu/~u0703432/>).

**Irregularity threshold test**

We extracted the irregularity traces of the right hand ascending scales for three pianists (CA, ES, and MD). The irregularity vector was multiplied by a scalar *factor* (between 0 and 5) and was then written as a MIDI file with eight notes per second, preceded by two metronome clicks at 120 BPM. For example, a factor of 0 means a perfectly regular (i.e., isochronous) scale, a factor of 1 corresponds to the scale as it was played in actuality, and a factor of 5 means that all note timings are five times more early or late than they were in reality whilst keeping the overall tempo intact. Participants were asked to report whether the scale sounded regular or irregular.

We used the maximum likelihood procedure (MLP) (Green, 1993; Gu and Green, 1994) to detect the threshold of the factor variable. Participants performed three thresholding blocks, one for each of the sample fingerprints. At the beginning of each block, we deployed 500 hypothetical psychometric curves with their midpoints linearly spaced over the factor levels from 0 to 5, crossed by the five false alarm rates of 0, 10, 20, 30, and 40%, yielding a total of 2,500 hypothetical psychometric curves maintained online in parallel. The slope parameter of these curves was set to four, since no prior experimental data exists and the slope has been shown not to influence the resulting thresholds all that much (Gu and Green, 1994). This yielded the following equation for the psychometric curves:  $p(\text{yes}) = a + (1 - a) \times (1 / (1 + \exp[-k \times (x - m)]))$ , where  $x$  is the stimulus level (i.e., the factor),  $a$  is the false alarm rate,  $m$  is the mean of the psychometric curve,  $k$  is the slope parameter (4),  $p(\text{yes})$  is the probability of responding “irregular.”

Each block consisted of 36 trials. On each trial, we calculated online the likelihood of the set of previous participant responses for each of the 2,500 hypothetical psychometric curves. The curve with the maximum likelihood was chosen as the current estimate. The magnification factor for that given trial was determined by the 64%-response point of this current estimate psychometric curve. In this way, the algorithm is shown to converge rapidly to the participant’s threshold (Green, 1993). We furthermore inserted two catch trials (with factor level 0 regardless of the

**Table 1 | Stimuli for the recognition experiment.**

Pianist (first)	Pianist (follow-up)	Comparison	Fingerprint distance (ms)	SD-IKI first (ms)	SD-IKI follow-up (ms)
CA	CA	Self	3.34	7.78	5.08
CA	MD	Other	7.28	7.78	6.65
ES	ES	Self	3.38	8.85	9.35
ES	IM	Other	8.30	8.85	8.39
TY	TY	Self	3.37	7.29	7.42
TY	VH	Other	7.34	7.29	9.22

SD-IKI is the Standard Deviation of the inter-keystroke-intervals (in ms).

current psychometric curve estimate) the first 12 trials at random locations, as well as four more over the remaining 24 trials.

Stimuli were written as MIDI files and then played through Timidity++ on a Windows computer, called by our Python (Pygame) graphical interface that registered the responses. The MLP computation was implemented in Python-MLP (which we have made available open-source online at: <https://github.com/florisvanvugt/PythonMLP>).

### Participants

Ten pianists from the Hanover University of Music student pool participated in this perceptual experiment. Participants (four female) were 24.8 (SD 3.7) years old and studied piano as their primary instrument. Further, they had normal hearing and reported no neurological impairments. The experiment took approximately half an hour and participants received a nominal payment for their participation.

## RESULTS

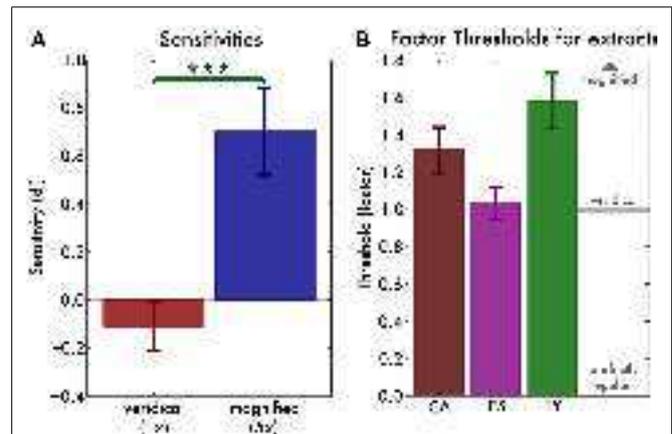
### Recognition test

We used signal detection theory to calculate sensitivity ( $d'$ ) for the individual participants, fingerprint pairs, and the factors (veridical or magnified) separately. There was a main effect of factor [ $F(1, 9) = 10.84, p = 0.001, \eta_G^2 = 0.25$ ], reflecting that sensitivity was greater for magnified (mean  $d' = 0.70$ , SD = 0.58) than for veridical (mean  $d' = -0.11$ , SD = 0.31) pairs (Figure 3A). There was no main effect of fingerprint pair [ $F(2, 18) = 1.44, p > 0.2$ ] but there was an interaction between factor and fingerprint pair [ $F(2, 18) = 6.09, p < 0.01, \eta_G^2 = 0.23$ ]. As a result, we investigated the sensitivity for each extract separately. For the veridical renditions, none of the sensitivities significantly exceeded zero [all  $t(9) < 0.7, p > 0.25$ ], indicating that participants were not able to distinguish pairs of recordings from the same pianist from pairs from different pianists. However, for the magnified renditions of the CA-MD and ES-IM pairs, sensitivity was significantly above zero [ $t(9) = 2.79, p = 0.01$ , and  $t(9) = 3.85, p < 0.01$ , respectively]. Only for the magnified TY-VH pair participants' sensitivity was zero [ $t(9) = 0.58, p = 0.29$ ].

After completing all blocks in this part of the experiment, participants were asked to subjectively rate the confidence in their answers on a five-point Likert scale from very confident (1) to very unsure (5). For the magnified fingerprint pairs, participants were mildly confident (median 3.5, range 2–4). For the veridical pairs, participants were similarly confident (median 4, range 3–5). The ratings did not differ significantly (Mann–Whitney  $U = 10.5, p = 0.29$ ). We can conclude that although participants performed much better in the magnified pairs, they were not aware of this improvement in performance.

### Irregularity threshold test

We discarded blocks in which participants' "irregular" response ratio for the catch trials exceeded 30%. This was the case for one block of one participant. The threshold for the remaining blocks was defined as the midpoint of the maximum likelihood estimate psychometric curve. Overall, curve midpoints expressed as factor were around or slightly above one (Figure 3B), meaning that the irregularities became audible only when they



**FIGURE 3 | (A)** Main effect of factor (veridical or magnified) in the recognition experiment. Sensitivity ( $d'$ ) is not greater than zero for the veridical rendering (factor 1), but is greater than zero for the magnified (factor 5) rendering. The error bars indicate the standard error of the mean. **(B)** Irregularity thresholds for three representative fingerprints. We find that the thresholds for all three extracts are one or above, that is, their irregularity is heard only when we exaggerate it slightly. Error bars indicate the standard error of the mean.

were slightly increased (factor  $>1$ ). The thresholds were entered into a one way ANOVA with fingerprint (the three example fingerprints) as a factor. There was a main effect of fingerprint [ $F(2, 26) = 4.85, p = 0.02, \eta_G^2 = 0.27$ ], indicating that the threshold factors were different for the different extracts. However, the fingerprints differed in evenness at the outset (see Table 1). As a result, we expressed the threshold not as a factor but as the corresponding unevenness value (SD of the inter-keystroke-intervals). We then re-ran the ANOVA and found no main effect of extract [ $F(2, 26) = 1.58, p = 0.22$ ]. The average threshold unevenness threshold value was 10.22 ms (SD 2.51).

## DISCUSSION

From our threshold experiment, we can conclude that the thresholds straddle the boundary of the timings as actually played (i.e., slightly above factor 1). Our interpretation is that pianists train to make their scale playing more regular until the irregularities are no longer audible.

We conclude that participants are not able to tell the difference between a scale as played by a pianist and an isochronous scale. It naturally follows that they will then not be able to differentiate between pianists since both scales sound regular (isochronous) to them. Indeed, in our recognition test participants were unable to distinguish pairs of scales played by the same pianist from pairs played by different pianists. However, when we magnified the timing deviations by a factor of five, the participants performed above chance in the recognition task. This shows that, in principle, the task of distinguishing scale playing of one pianist from another can be done. These two tests, taken together, constitute evidence that participants were not able to hear the differences between the pianist fingerprints and categorize them on the basis of these differences.

Our study is also the first to systematically investigate thresholds for perception of irregularity in piano scales. We find that the irregularities in recorded piano scales are slightly below the perceptual threshold. This in itself is an interesting finding. Our interpretation is that pianists practice to make their scale playing sound regular but do not continue to make it more regular once it is below the perceptual threshold. For one, listeners will not be able to tell the difference, and secondly, if the motor learning of scale regularity is guided by auditory feedback (Jäncke, 2012) only, they will not be able to improve their temporal regularity once they fall below the auditory threshold.

We furthermore found that the differences in threshold between the extracts can be explained by their difference in unevenness: more temporally uneven fingerprints have a lower factor threshold, whereas more temporally even fingerprints have a higher threshold. This suggests that the obtained threshold of 10.22 ms is independent of the particular temporal fingerprint. We conclude that the unevenness captures the auditory percept of unevenness and no more complex auditory gestalt needs to be taken into account to explain the thresholds. The threshold corresponds to some 8.2% of the interval at this tempo, which is in line with the typical 10% threshold of a single late or early note in an otherwise isochronous sequence (Hyde and Peretz, 2004; Ehrle and Samson, 2005).

Since these individual characteristics of the scale fingerprints are inaudible, it seems that their production is not dependent on auditory feedback. However, this conclusion is not warranted, since it could be that the timing deviations are residuals of expressive timing (Repp, 1999a). To clarify this issue, we investigated whether the pianistic fingerprints were affected by playing on a mute piano.

### EXPERIMENT III

#### MATERIALS AND METHODS

Eighteen piano students (nine female) from the Hanover University of Music were invited to play two-octave C-major scales in two recordings. Participants were 28.2 (SD 5.8) years old. In the first recording, participants heard the sounds they produced (*sound*) but in the second recording the sound was switched off (*mute*). In both recordings, scales were played by one hand and then by the other. Otherwise, the procedure and analysis was identical to before. We report 95% confidence intervals (CI) unless otherwise stated.

#### RESULTS

We discarded incorrectly played scales leaving a total of 13.4 (SD 1.77) per condition. There was no effect of hand, direction, or recording on the number of correctly produced scales [all  $F(1, 17) < 1.8$ ]. There was a significant but marginally small interaction between hand and direction [ $F(1, 17) = 4.71$ ,  $p = 0.04$ ,  $\eta_G^2 = 0.001$ ] and none of the other interactions was significant [all  $F(1, 17) < 4.3$ ].

As before, the distances between fingerprints originating from the same pianist are smaller than those originating from different pianists [ $F(1, 17) = 168.2$ ,  $p < 0.001$ ,  $\eta_G^2 = 0.55$ ]. There was a (small) interaction between hand and direction [ $F(1, 17) = 7.45$ ,  $p = 0.01$ ,  $\eta_G^2 = 0.03$ ], indicating that for the right hand, inward

scales are more similar than outward scales, whereas for the left hand this was the opposite.

Our distance-minimizing algorithm introduced in Experiment I correctly recognized between 8 (44%) and 12 (67%) of the 18 pianists using the fingerprint for only one hand and direction at a time. This exceeds chance performance, which lies at 6%. The correlation-maximizing algorithm correctly recognized between 7 (39%) and 15 (83%) pianists.

When we combined the two hands and two directions (yielding a  $2 \times 2 \times 15$  fingerprint matrix for each participant) and perform the same classification, the distance-minimizing algorithm correctly identified 15 out of 18 pianists [83%, binomial  $p < 0.001$ , confidence interval (0.59, 0.96)]. Crucially, the result is the same whether matching the *mute* fingerprints, one by one, to the set of *sound* fingerprints, or the other way around, indicating that there is no loss of information in the *mute* condition. The correlation-maximizing algorithm also recognizes 15 out of 18 pianists when it finds matching *sound* fingerprints to a given *mute* fingerprint, and the other way around spectacularly recognizes all 18 pianists [100%, binomial  $p < 0.001$ , CI (0.81, 1.00)].

In order to compare our results with those of Experiment I, we take 10,000 bootstrap samples of eight (unique) pianists and perform the classification with those. The correlation-maximizing algorithm recognizes 95% of pianists [SD 8%, bootstrap CI (75, 100)] whereas the distance-minimizing algorithm recognizes 90% of pianists [SD 8%, bootstrap CI (75, 100)]. That is, they do not perform significantly differently.

#### DISCUSSION

It is becoming clear that having auditory feedback while playing the scales is not of importance in the formation of the pianistic fingerprint. Indeed, it is a typical finding in performance literature that absence of auditory feedback only marginally affects performance (Repp, 1999b) or not at all (Gates and Bradshaw, 1974). The findings are furthermore in line with our previous result that fingerprints generated by the same movements are more similar than those generating the same sounds (Experiment I).

Finally, we turn to the question of how stable these fingerprints are over time.

### EXPERIMENT IV

#### MATERIALS AND METHODS

We re-analyzed data published previously (Jabusch et al., 2009) in which 20 pianists' (eight female) scale playing was measured twice (first, follow-up) with an interval of 27.8 (SD 8.8) months. At the first measurement, pianists were 27.7 (SD 6.0) years old and had accumulated 21.6 (SD 11.0) thousand hours of lifetime piano practice (not counting one pianist who had not reliably reported this figure). In between the two measurement sessions, pianists accumulated an additional 2.8 (SD 1.8) thousand practice hours, amounting to an average 3.31 (SD 1.79) hours per calendar day (including weekends and holidays). All but two pianists were right-handed according to the Edinburgh handedness inventory (Laterality Quotient:  $M = 73\%$ , SD 56).

#### RESULTS

After discarding incorrect scales we were left with 13.5 (SD 0.8) scales of the *first* measurement and 12.8 (SD 1.2) scales at

the *follow-up* measurement. This difference was significant [ $F(1, 19) = 5.65$ ,  $p = 0.03$ ,  $\eta_G^2 = 0.09$ ].

As before, distance was smaller between recordings of the same pianist than that of different pianists [ $F(1, 19) = 184.90$ ,  $p < 0.001$ ,  $\eta_G^2 = 0.30$ ]. Furthermore, distance was generally smaller between fingerprints of the right hand than those of the left hand [ $F(1, 19) = 6.33$ ,  $p = 0.02$ ,  $\eta_G^2 = 0.05$ ], perhaps reflecting the greater training of the right hand (Kopiecz et al., 2011). For brevity, we only report recognition results using the fingerprint combining both hands and directions. Recognition based on minimizing distance successfully found first recordings given the follow-up fingerprints in 13 pianists [65%, binomial  $p < 0.001$ , CI (40, 85)%]. Conversely, seven pianists were recognized based on their follow-up measurement [35%, binomial  $p < 0.001$ , CI (15, 59)%]. Recognition by maximizing correlation performed similarly with 13 (65%) and 8 [40%, binomial  $p < 0.001$ , CI (19, 64)%] correct identifications.

Bootstrap analysis was performed (see Experiment III) with 10,000 samples of eight pianists. Correlation recognition identified 73% [SD 15%, bootstrap CI (38, 100)%] of pianists and distance recognition 71% [SD 18%, bootstrap CI (38, 100)%]. Based on the bootstrap CI we can see that across the three experiments, identification was equally successful.

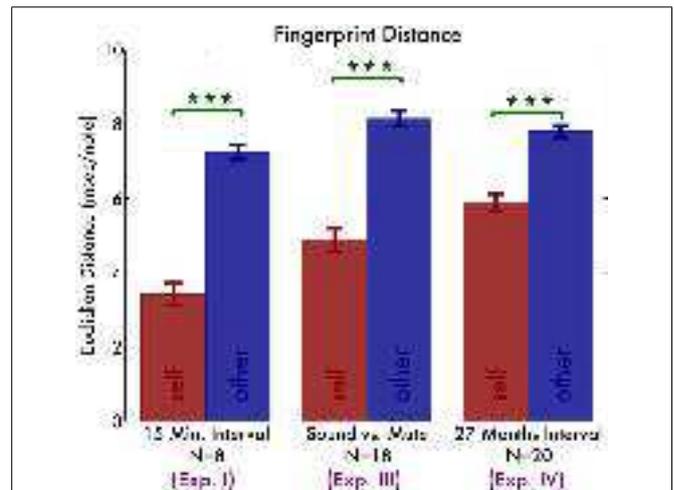
How is the stability of a pianist's fingerprint related to how much he or she practised between the two measurements? We calculated the distance for both hands and playing directions and correlated this to the number of practice hours accumulated between the two measurement points. The distances between the right hand outward scale fingerprints correlated negatively with amount of practice (Pearson  $r = -0.71$ ,  $p = 0.001$ ). That is, those who practised more showed smaller distances between their fingerprints. This does not mean that the fingerprints showed less deviations from regularity, but instead, that the deviations that were present were more consistently reproduced. The right hand inward fingerprints showed a tendency for the same correlation (Pearson  $r = -0.46$ ,  $p = 0.05$ ) but the left hand fingerprints did not ( $r > -0.34$ ,  $p > 0.16$ ).

**DISCUSSION**

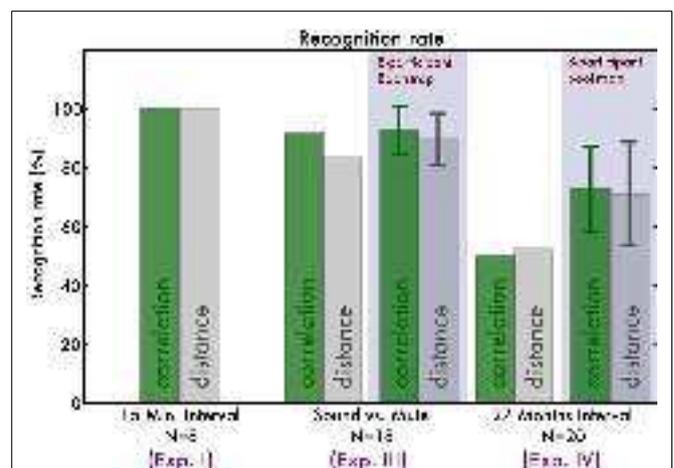
The fingerprints that enabled reliable identification of pianists were sufficiently stable to still allow recognition after 27 months. **Figure 4** compares the distances across the Experiment I, III, and IV and **Figure 5** displays the recognition rates. Although it seems the recognition is worse in Experiment III and IV, the 95% bootstrap CI still include the 100% recognition rate of Experiment I. Therefore we conclude that recognition is not significantly different across the experiments.

**GENERAL DISCUSSION**

Artists are recognized reliably based on their work (Yamamura et al., 2009). The present study investigated pianist recognition based on non-expressive materials. Taking scale playing as an example, this study brings to light a highly individual temporal signature that enables robust identification of pianists using a simple algorithm. Clearly an individual timing signature is present physically, but perceptual recognition performance by musician listeners was at chance because the deviations were below



**FIGURE 4 | Summary of the distances between fingerprints originating from the same pianist (self; the red bars) and fingerprints originating from different pianists (other; the blue bars).**



**FIGURE 5 | Overview of the recognition rates of our recognition algorithm.** The green bars indicate the correct classification rate by maximizing fingerprint correlation, and the gray bars by minimizing fingerprint distance. For comparison, we indicate the bootstrap classification results, indicating for each experiment the average recognition rates across eight-pianist bootstrap samples. Error bars indicate the standard deviation of the recognition rates.

their perceptual thresholds. Fingerprints appear to stem from neuromuscular factors in the pianists, rather than auditory feedback. This is confirmed in Experiment III that shows fingerprint formation is not affected by absence of sound. The fingerprint is furthermore robust, showing only mild changes in professional pianists over a 27-month interval.

The findings are in line with previous studies showing that pianists can be reliably recognized even when asked to not play expressively (Gingras et al., 2011). Our result strengthens the interpretation that recognition is based on non-expressive clues by employing materials (musical scales) with a clear auditory

goal of regularity. Moreover, we have at present only used timing information, discarding loudness and articulation markings that could potentially be used to enhance recognition. The recognition algorithm that we present pairs fingerprints with minimum distance or maximum correlation. The proposed similarity metric is transparent and easy to interpret (see **Figure 1**). As such, it is surprisingly simple compared with neural networks typically employed (Stamatatos and Widmer, 2005; Dalla Bella and Palmer, 2011).

The idea that artists can be recognized by a non-artistic feature of their work is not new. For example, painters can be automatically recognized by stroke style (Li et al., 2012). Beyond the realm of art, authorship can be established by relatively irrelevant features of produced work. For example, handwriting is highly individual (Rijntjes et al., 1999) and pattern recognition using word frequencies has been employed to establish Madison as the author of the 12 disputed Federalist papers (Mosteller and Wallace, 1964). Similarly, telegraph operators during the Second World War claimed to be able to identify the sender by the timing of his keystrokes (“Fist of sender”). The emerging field of keystroke dynamics puts this to use to authenticate computer users by their typing rhythm instead of through a password (Bergadano et al., 2002). Typically the problem remains that over time these dynamics change and recognition becomes impaired. In light of this, it is interesting that our recognition was highly stable even in a fairly homogeneous sample of expert pianists (Experiment IV). Recognition in keystroke dynamics as well as in our result may be based to some extent on the subunits that the produced sequences are divided

into, i.e., its chunking (Sakai et al., 2003). On the other hand, more low-level neuromuscular properties such as the individual anatomy, especially tendon-ligament anatomy or the strengths of the individual muscles are more likely to be at the root of these individual temporal irregularities, since the sequences under consideration here (the scales) are greatly over-learned. Future studies may decide this issue by investigating recognition of pianists playing at various tempi, since although chunking may vary across speeds, the neuromuscular properties will remain constant.

We propose that studies investigating the individuality of artists, especially those employing machine learning strategies (Stamatatos and Widmer, 2005), may take into account that a large part of this individuality is inaudible and merely neuromuscular in nature. In the future, one could tease apart cues that are uniquely expressive and those that are neuromuscular.

Artistic individuality is typically thought to be deliberate and determined by top-down cognition. Our study opens the road to investigation into the tantalizing question of how biomechanical constraints may determine artistic performance in a bottom-up fashion.

## ACKNOWLEDGMENTS

This work was supported by the EBRAMUS, European Brain and Music Ph.D. Grant to Floris Tijmen Van Vugt (ITN MC FP7, GA 238157). The authors are indebted to Laura Lübke data collection in Experiment II, as well as Melinda Lee for running Experiment III. We also thank Meaghan Fowlie for proofreading the manuscript.

## REFERENCES

- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384.
- Bergadano, F., Gunetti, D., and Picardi, C. (2002). User authentication through keystroke dynamics. *ACM Trans. Info. Syst. Secur.* 5, 367–397.
- Dalla Bella, S., and Palmer, C. (2011). Rate effects on timing, key velocity, and finger kinematics in piano performance. *PLoS ONE* 6:e20518. doi:10.1371/journal.pone.0020518
- Drake, C. (1993). Perceptual and performed accents in musical sequences. *Bull. Psychon. Soc.* 31, 107–110.
- Ehrlé, N., and Samson, S. (2005). Auditory discrimination of anisochrony: influence of the tempo and musical backgrounds of listeners. *Brain Cogn.* 58, 133–147.
- Engel, K. C., Flanders, M., and Soechting, J. F. (1997). Anticipatory and sequential motor control in piano playing. *Exp. Brain Res.* 113, 189–199.
- Flach, R., Knoblich, G., and Prinz, W. (2004). Recognizing one’s own clapping: the role of temporal cues. *Psychol. Res.* 69, 147–156.
- Gates, A., and Bradshaw, J. L. (1974). Effects of auditory feedback on a musical performance task. *Percept. Psychophys.* 16, 105–109.
- Gingras, B., Lagrandeur-Ponce, T., Giordano, B. L., and McAdams, S. (2011). Perceiving musical individuality: performer identification is dependent on performer expertise and expressiveness, but not on listener expertise. *Perception* 40, 1206–1220.
- Green, D. M. (1993). A maximum-likelihood method for estimating thresholds in a yes–no task. *J. Acoust. Soc. Am.* 93, 2096.
- Gu, X., and Green, D. M. (1994). Further studies of a maximum-likelihood yes–no procedure. *J. Acoust. Soc. Am.* 96, 93.
- Hill, H., and Pollick, F. E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychol. Sci.* 11, 223–228.
- Hyde, K. L., and Peretz, I. (2004). Brains that are out of tune but in time. *Psychol. Sci.* 15, 356–360.
- Jabusch, H.-C., Alpers, H., Kopiecz, R., Vauth, H., and Altenmüller, E. (2009). The influence of practice on the development of motor skills in pianists: a longitudinal study in a selected motor task. *Hum. Mov. Sci.* 28, 74–84.
- Jabusch, H.-C., Vauth, H., and Altenmüller, E. (2004). Quantification of focal dystonia in pianists using scale analysis. *Mov. Disord.* 19, 171–180.
- Jäncke, L. (2012). The dynamic audiomotor system in pianists. *Ann. N. Y. Acad. Sci.* 1252, 246–252.
- Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behav. Brain Res.* 142, 1–15.
- Kopiecz, R., Jabusch, H.-C., Galley, N., Homann, J.-C., Lehmann, A. C., and Altenmüller, E. (2011). No disadvantage for left-handed musicians: the relationship between handedness, perceived constraints and performance-related skills in string players and pianists. *Psychol. Music* 40, 357–384.
- Li, J., Yao, L., Hendriks, E., and Wang, J. Z. (2012). Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 1159–1176.
- Loula, F., Prasad, S., Harber, K., and Shiffrar, M. (2005). Recognizing people from their movement. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 210–220.
- MacKenzie, C. L., and Van Eerd, D. L. (1990). Rhythmic precision in the performance of piano scales: motor psychophysics and motor programming. *Atten. Perform.* 13, 375–408.
- Mosteller, F., and Wallace, D. L. (1964). *Inference and Disputed Authorship: The Federalist*. Reading: Addison-Wesley.
- Murgia, M., Hohmann, T., Galmonte, A., Raab, M., and Agostini, T. (2012). Recognising one’s own motor actions through sound: the role of temporal factors. *Perception* 41, 976–987.
- Prasad, S., and Shiffrar, M. (2009). Viewpoint and the recognition of people from their movements. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 39–49.
- Repp, B. H. (1999a). Control of expressive and metronomic timing in pianists. *J. Mot. Behav.* 31, 145–164.

- Repp, B. H. (1999b). Effects of auditory feedback deprivation on expressive piano performance. *Music Percept.* 16, 409–438.
- Repp, B. H., and Knoblich, G. (2004). Perceiving action identity: how pianists recognize their own performances. *Psychol. Sci.* 15, 604–609.
- Rijntjes, M., Dettmers, C., Büchel, C., Kiebel, S., Frackowiak, R. S. J., and Weiller, C. (1999). A blueprint for movement: functional and anatomical representations in the human motor system. *J. Neurosci.* 19, 8043–8048.
- Sakai, K., Kitaguchi, K., and Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Exp. Brain Res.* 152, 229–242.
- Sevdalis, V., and Keller, P. E. (2011). Perceiving performer identity and intended expression intensity in point-light displays of dance. *Psychol. Res.* 75, 423–434.
- Stamatatos, E., and Widmer, G. (2005). Automatic identification of music performers with learning ensembles. *Artif. Intell.* 165, 37–56.
- van Vugt, F. T., Jabusch, H.-C., and Altenmüller, E. (2012). Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Front. Psychol.* 3:495. doi:10.3389/fpsyg.2012.00495
- Wagner, C. (1971). “The influence of the tempo of playing on the rhythmic structure studied at pianist’s playing scales,” in *Medicine and Sport Biomechanics II*, eds J. Vredenburg and J. Wartenweiler (Basel: Karger), 129–132.
- Yamamura, H., Sawahata, Y., Yamamoto, M., and Kamitani, Y. (2009). Neural art appraisal of painter: Dali or Picasso? *Neuroreport* 20, 1630–1633.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 24 January 2013; paper pending published: 12 February 2013; accepted: 04 March 2013; published online: 21 March 2013.
- Citation: Van Vugt FT, Jabusch H-C and Altenmüller E (2013) Individuality that is unheard of: systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint. *Front. Psychol.* 4:134. doi: 10.3389/fpsyg.2013.00134
- This article was submitted to *Frontiers in Cognitive Science*, a specialty of *Frontiers in Psychology*.
- Copyright © 2013 Van Vugt, Jabusch and Altenmüller. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



*Illustration 11: Artistic impression of the pianistic fingerprint. Drawing by Vera Scepanovic, based on an idea by Mariana Fulgueiras Montoro.*

**IV.4. Spatial and temporal symmetries of motor primitives in  
skilled piano performance at different tempi.**

---

## **Spatial and temporal symmetries of movement patterns in skilled piano performance at different tempi**

Floris T. van Vugt<sup>1</sup>, Shinichi Furuya<sup>1</sup>, Henning Vauth<sup>1</sup>, Hans-Christian Jabusch<sup>2</sup>, Eckart Altenmüller<sup>1</sup>

<sup>1</sup>University of Music, Drama, and Media, Hanover, Germany

<sup>2</sup>Dresden University of Music "Carl Maria von Weber", Dresden, Germany

### **Abstract**

In order to understand the processes that allow motor actions to be performed at various speeds, we investigated pianists playing scales at a wide range of tempi. Systematic timing deviations followed two qualitatively different patterns. One pattern was essentially a slowing down at the end of the scale and this pattern dominated at slow tempi. The second pattern was a series of three peak delays caused by a thumb-under manoeuvre and this pattern dominated the fast tempi. Intermediate tempi were a linear combination trade-off between these two. We fit a simple five-parameter model that predicts the timing deviation of each individual note across tempi and captures a large part of the variation ( $R^2=0.70$ ) in this highly complex data set. We then show that the motor system is able to generalise these two motor primitives across the effectors (hands) and by movement directions. In sum, the motor system implements a dimensionality reduction in performance at varying speed demands through a dynamic combination of motor primitives that can be inverted in time and space.

Keywords: piano scales, generalised motor programs, dimensionality reduction, spatial symmetry, temporal symmetry

## 4.1 Introduction

Sequential motor behaviour such as speech, finger spelling, typing a text, and music performance is a way of conveying information, a crucial aspect of human communication. In many cases, the sequential behaviours exhibit spontaneous temporal unevenness as a result of cognitive, perceptual, emotional and biomechanical factors. For example, music performance reveals some systematic patterns of temporal deviation from an evenly timed succession of events (Palmer, 1997; Repp, 1990, 1992; van Vugt, Jabusch, & Altenmüller, 2012). A key issue is how the nervous system scales the spatio-temporal patterns of the sequential movements in various contexts (e.g. varying speed, force, or direction), since spatio-temporal patterns are a means to transmit a variety of information content (Angelaki & Soechting, 1993; d' Avella, Portone, Fernandez, & Lacquaniti, 2006; Lacquaniti, Ferrigno, Pedotti, Soechting, & Terzuolo, 1987). Is the pattern of temporal unevenness different across different movement contexts? Or is there a set of basic movement patterns, so-called motor primitives, serving as building blocks of production of rich repertoires of motor behaviours? Understanding this may provide important insights into principles governing production of diverse movement behaviours at different contexts.

Playing a musical sequence at various global and local tempi conveys different emotional information to listeners (Bhatara, Tirovolas, Duan, Levy, & Levitin, 2011; Dalla Bella, Peretz, Rousseau, & Gosselin, 2001; Khalfa, Roy, Rainville, Dalla Bella, & Peretz, 2008). The sensorimotor skill that enables musicians to manipulate tempo thus plays a crucial role in expressive musical performance. While previous studies focused on variation of the organization of movement kinematics and muscular activity in relation to tempo during piano playing (Furuya, Aoki, Nakahara, & Kinoshita, 2012; Furuya, Goda, Katayose, Miwa, & Nagata, 2011), it is not well understood how temporal features of successive movement elements (i.e. keystrokes) vary across a wide range of tempi.

On the one hand, movement production over a wide range of tempi can emerge as a consequence of parametric modulation of a single temporal movement template or generalised motor program (Shapiro, Zernicke, & Gregor, 1981). For example, when participants learn a tapping sequence in a certain rhythm, this rhythm is preserved even

when they are asked to reproduce the sequence as fast as possible (Summers, 1975). Such a tempo-invariant motor program simplifies players' manipulation of tempo of movements, and predicts invariant temporal features of successive keystrokes across tempi in piano playing. This prediction was supported by an exploratory study that reported maintained temporal features of successive keystrokes across three different tempi (Repp, 1994). Similarly, in discrete motor tasks such as reaching (Hollerbach & Flash, 1982) and ball-throwing with the non-dominant hand (Hore, O'Brien, & Watts, 2005), spatio-temporal features of movements were maintained across different speeds.

However, an alternative hypothesis proposes that variation of tempo in music performance involves combining a particular set of distinct movement templates. This may be computationally costly, but makes one flexible enough to cope with cognitive and biomechanical demands that can vary in relation to tempo. For example, disruption of movements by altered auditory feedback in piano playing differed according to tempo, implying differential reliance on feedback control across tempi (Furuya & Soechting, 2010). In addition, when playing some types of musical scales, the thumb has to rotate while pressing a key in order to move the hand horizontally (Engel, Flanders, & Soechting, 1997; Furuya, Flanders, & Soechting, 2011), which would biomechanically constrain keystrokes in terms of the hand inertia force that changes with tempo. The tempo-modulation strategy using a linear sum of some fundamental movement patterns predicts variation of movement characteristics (or more specifically, systematic pattern of spontaneous temporal deviation) with tempi. Some previous studies supported this hypothesis by demonstrating violation of temporal invariance of successive keystrokes across tempi in piano performance (Clarke, 1982; Desain & Honing, 1994) as well as in particular for scale playing (MacKenzie & Van Eerd, 1990; Wagner, 1971).

Independent of the question of whether motor actions at various speeds are controlled by a single or by multiple movement templates, is the question in what ways these motor primitives can be recombined. How specific are these motor primitives to the effector (e.g. hand) in question? For example, because of better control of non-muscular force such as inter-segmental dynamics at the dominant hand compared to non-dominant hand (Heuer, 2007; Sainburg, 2002; Sainburg & Kalakanis, 2000) we may predict distinct movement control between the two hands. On the other hand, it may be computationally

advantageous to adaptively use effector-shared motor representations when both hands execute the same movement.

A second question is how sensitive the motor primitives are to the temporal direction in which a movement is executed? That is, could the same motor primitive govern a movement that is executed backwards in time? To our knowledge, this question of directionality in motor representations has not been addressed before. Executing a movement sequence in a different order might introduce different muscle synergies, and therefore the movement would need to be represented at a higher level that abstracts away from the immediate muscular instruction to allow for such a generalisation across movement directions.

To test these questions straightforwardly, the present study developed a linear regression model of timing deviations in piano playing across various tempi, movement directions and effectors (the two hands). This enabled us to evaluate whether variability across successive keystrokes can be best accounted for in terms of one or multiple movement templates. To this aim, we asked highly skilled pianists to play musical scales over a wide range of tempi and assessed the effect of tempo on the temporal characteristics of successive keystroke motions quantitatively. We used musical scales since they are basic elements of the musical architecture in classical music as well as in jazz, rock, and pop music. As a consequence, scale playing is a fundamental aspect of piano technique. With respect to the present study, scales represent a model, which allows us to investigate the influence of pianistic expertise, movement direction, and speed on movement characteristics. We specifically tested the following hypotheses. First, we tested the hypothesis that variation of tempo in scale playing is a parametric modulation of a single generalised motor pattern. Predictions of this hypothesis are a monotonic increase in temporal unevenness and loudness with increasing tempo. Alternatively, there may be multiple (two or more) motor templates at play. In this case, we set out to investigate how the trade-off between these representations is controlled across tempi. Thirdly, we hypothesise that these representations can be flexibly mirrored in space and, perhaps to a lesser extent, in time. Finally, we expect that more advanced pianists compensate more for biomechanical constraints in scale playing.

## 4.2 Methods

### 4.2.a Participants

Nineteen piano students (ten female, nine male), aged on average 23.7 (SD 2.9) years, were recruited from the student pool at the Hanover Music University and recorded playing two-octave C major scales. All but one were right-handed and except for two organ students all took piano as their main subject. The pianists had no history of neurological disorders and reported an accumulated lifetime keyboard practice of 16.9 (SD 7.5) thousand hours (except for two pianists who did not specify). The experiment was performed in accordance with the Declaration of Helsinki.

### 4.2.b Procedure

Participants were asked to play two-octave C-major scales, beginning with the C (131 Hz) that is located one octave lower than the middle C and ending with the C (523 Hz) that is located one octave higher than the middle C (ascending) and back (descending). For notational convenience, we will write the notes in this scale as *c, d, e, f, g, a, b, c', d', ..., c''*. To play a two-octave scale, the participants needed to press fifteen adjacent keys serially. Playing the ascending scale with the right hand, the thumb starts out on the *c*, then the index and middle finger press the adjacent *d* and *e*. Now the thumb has to move underneath the index and middle finger to strike the next key, *f*. The role of this manoeuvre is to move the position of the hand rightward. Now the index finger moves over the thumb to strike the *g*. The middle and ring finger can then in turn strike the *a* and *b*. At this point, another thumb-under movement is performed to strike the *c*, at which point the entire movement is repeated for the next octave. The left hand movements are analogue for the descending scale starting with the thumb on *c''*. In sum, participants played the scales with the conventional fingering (123123412312345 and reverse, where the fingers are numbered from 1=thumb to 5=little finger).

In between the ascending and descending scales there was a short pause of about a

second and the pianists made one additional keystroke on either the note C (65 Hz) or C (1047 Hz) with the hand not involved in the scale. This keystroke was used in previous analyses to demarcate the different trial runs, but was discarded in our current analysis. The procedure was repeated for 5 different tempi indicated by a metronome at either 40, 80, 120, 160 or 200 BPM. The procedure was then repeated for the other hand. The pianists were instructed to play 4 keystrokes per metronome beat, which means that the keystroke rate varied between 2.67 and 13.3 keystrokes/sec. They were instructed to aim for temporal evenness.

Participants played on a MP 9000 MIDI keyboard (Kawai, Krefeld, Germany). The keyboard's digital music interface (MIDI out) signal was captured on a PC using a commercially available sequencer software (Musicator Win, version 2.12; Music Interactive Technology, Bergen, Norway).

#### **4.2.c Processing of MIDI data**

First, we extracted only the correctly played scales, separating ascending and descending scales. This yielded 13.5 (SD 1.2) scales for each pianist, tempo, hand and direction. Across participants a total of 1615 keystroke onsets were discarded because they were part of scales with omissions or wrong notes, amounting to 2.1% of the recorded material. For the individual pianists, this value ranged from 0.8% to 10.9%. The average amount of dropped onsets was 3.2% (SD 2.6%). This amount of discarded note material did not correlate with the accumulated practice hours of the pianist (taken as a measure of pianistic expertise) (Spearman  $\rho(14)=-.16$ ,  $p=.57$ ).

#### **4.2.d Scale unevenness**

Firstly we reproduced the analysis of keystroke unevenness as introduced previously (Jabusch, Vauth, & Altenmüller, 2004; Wagner, 1971). For each correctly played scale, we computed the intervals between the keystroke onsets and then the standard deviation of these intervals. This gave a single “unevenness” value for each scale run (in msec). A higher unevenness value means that playing was more temporally variable, that is, less precise. A smaller unevenness value means less variable playing. We took the median

of this unevenness for each pianist and each condition (i.e., each tempo, hand, playing direction). The choice to compute the median instead of mean in this last step was due to the median's robustness to outliers.

#### 4.2.e Unevenness across tempi

At this point, we establish the dependency of unevenness across the various tempi. We proceed to test the generalised motor program prediction that unevenness increases (linearly) with tempo. Alternatively, if multiple motor representations are at play, we expect the unevenness curve to follow an alternative shape. We contrast a polynomial model that essentially predicts an optimal tempo with a minimum unevenness, with a two-process model that captures two processes, one of which dominates at the slow tempi and another that dominates at the fast tempi.

#### 4.2.f Individual note timing

The unevenness metric described above provides insight into the unevenness at the level of the entire scale. However, our aim was to further understand the timing deviations at the level of the individual notes within the scale. In order to establish the timing of individual notes we introduced the following procedure (Van Vugt et al., 2012): For each ascending and descending scale we computed the (note, time) pairs, where note is the rank of the note in the particular scale (0 for *c*, 1 for *d*, etc., until 14 for *c''*) and *time* represents the keystroke onset time. Subsequently, we fitted an ordinary least square regression line to these pairs. We rejected fits with an  $R^2$  of less than .9 (which occurred when a scale was played very unevenly), and this was the case for one scale (.02% of the data). The remaining fits had an average  $R^2$  of 1.0 ( $\pm 3 \cdot 10^{-4}$ ). The fitted line enabled us to determine the expected onset time for each keystroke as the intersection of the regression fit with the horizontal line representing the corresponding note. Finally, we computed the time difference between this expected onset and the actual keystroke onset (in msec, with negative difference reflecting notes played earlier than predicted).

#### 4.2.g Towards a model of timing deviations

Using the above procedure, we computed the median timing deviation for each note, hand, playing direction (inward and outward; inward being defined as radial playing direction, with the arm moving toward the body; outward as ulnar playing direction, with the arm moving away from the body) and tempo across players. At this point, we investigated whether the timing deviation traces for the various tempi were dominated by a single template or by multiple templates. We hypothesised to reproduce previous findings of movement template that reflect phrasing: a slowing down in the beginning and at the end of the movement. Additionally, to cope with changing demands in the context of higher movement production speed, the traces may be qualitatively different. We would expect such a template to be determined by the thumb- and index finger manoeuvre. We generated two timing templates that were represented as a vector of timing deviations for each note. The timing deviations at the intermediate production speeds were then rewritten as a linear combination of these two vectors.

Next, we provided a stepwise generalisation of our model across the hands and movement directions. First, we cross-validated the model fit on the right hand outward scales by assessing its fit on the left-hand outward scales, which are produced by the same movement but mirrored in space (using the symmetry axis of the body). Second, we turned to the inward scales for both hands, which are the same movement but mirrored in time (that is, played backwards). That is, we investigated whether the movement templates that we identified in the steps above are also mirrored in time (that is, played backwards) or whether they remain time-invariant.

#### 4.2.h Fitting methods

For fitting, we use the analytic ordinary least squares solution provided by ordinary least squares fitting for generalised linear models. For non-linear models that do not allow this, we instead use the Newton-type optimisation using R's *nlm* function. Furthermore, the non-linear models with constrained parameter spaces were implemented using R's *optim* function using the Nelder-Mead simplex method.

In ANOVAs we report  $\eta_G^2$  as the generalised effect size (Bakeman, 2005). Whenever Mauchly's test indicated sphericity violations, we applied the Greenhouse-Geisser

correction. In those cases, for the sake of brevity, we only report the p-value after the correction is applied and marked it as  $p_{GG}$ . Whenever appropriate, we report adjusted goodness-of-fit statistics  $R^2_{adj} = 1 - ((1 - R^2) * (N - 1) / (N - k - 1))$  where  $R^2$  is the conventional goodness-of-fit,  $N$  is the number of observations and  $k$  is the number of parameters in the model. Alternatively, whenever the variance of the observed data points was important, we reported the chi-square measure of goodness-of-fit as follows:  $\chi^2_{red} = 1 / (N - k - 1) * (\sum_i (\text{Observed}_i - \text{Predicted}_i)^2 / \text{Variance}_i)$ , where  $i$  is the index of the observation,  $N$  is the number of observations and  $k$  is the number of parameters in the model.

## 4.3 Results

### 4.3.a Scale-level tempo dependencies

We attempted to begin with testing whether a linear increase in unevenness obtains as a function of tempo, as predicted by the generalised motor program account. We submitted the evenness measure to a three-way repeated-measures ANOVA using 5 tempi x 2 hands x 2 playing directions (inward/outward) as factors. To balance the ANOVA, we eliminated the two organists who could not play the fastest tempo. The direction of the scales (ascending or descending) was translated to body-central direction (outward or inward), since timing deviations were more similar between the same (mirrored) movements (e.g., right hand ascending and left hand descending) than between the same keystroke series (e.g., right hand ascending and left hand ascending) (Van Vugt, Jabusch, & Altenmüller, in revision).

We found a main effect of hand ( $F(1,16)=14.70$ ,  $p<.01$ ,  $\eta_G^2=.03$ ), revealing that the right hand is more temporally precise ( $M=11.9$   $SD=1.07$  msec) than the left hand ( $M=12.7$   $SD=1.13$  msec). This effect has previously been described (Kopiez et al., 2011). There was furthermore a main effect of direction ( $F(1,16)=9.03$ ,  $p<.01$ ,  $\eta_G^2=.02$ ), reflecting that outward scales were more uneven ( $M=12.7$   $SD=.99$  msec) than inward scales ( $M=12.0$   $SD=1.20$  msec). There was also a main effect of playing tempo ( $F(4,64)=32.85$ ,  $p_{GG}<10^{-9}$ ,  $\eta_G^2=.47$ ). There were two-way interactions between tempo and hand ( $F(4,64)=5.28$ ,  $p_{GG}<.01$ ,  $\eta_G^2=.03$ ) but not

tempo and direction ( $F(4,64)=2.70$ ,  $p_{GG}>.06$ ) nor between hand and direction ( $F(4,16)=0.00$ ,  $p>.95$ ). Furthermore, the three-way interaction hand-tempo-direction was significant ( $F(4,64)=3.54$ ,  $p_{GG}=.02$ ,  $\eta_G^2=.01$ ). We therefore continued to analyse each of combinations of hand and direction separately, focusing first on the right hand outward scales.

Instead of a linear increase in unevenness, we found a u-shaped curve (Fig. 1a) with the lowest unevenness at the intermediate tempo of around 8 keystrokes/sec. This is in keeping with previous reports (Wagner, 1971). We then modelled the playing unevenness as a function of tempo. A simple quadratic polynomial fit ( $s=at^2+bt+c$  where  $s$  is the unevenness and  $t$  is the tempo) on this quantity yielded a moderate fit ( $R^2_{adj}=0.54$ ,  $F(2,90)=53.03$ ,  $p<10^{-15}$ ) but a fit of a two-term model  $s=at^{-1}+bt$  yields a much better fit ( $R^2_{adj}=0.96$ ,  $F(2,83)=1230$ ,  $p<10^{-15}$ ) which is significantly better ( $F(1,83)=4.68$ ,  $p=.03$ ). Now we turn to the other combinations of hand and direction. In the other combinations of hand and direction (data not plotted), the quadratic polynomial fit was significant (all  $R^2_{adj}>.40$ ,  $F(2,90)>27.20$ ,  $p<10^{-9}$ ) as was the two-term model (all  $R^2_{adj}>.96$ ,  $F(2,83)>1030.86$ ,  $p<10^{-15}$ ), and the latter was always better than the former (all  $F(1,83)>4.43$ ,  $p<.04$ ).

The finding of the superiority of our two-term model is interesting in our context, since the model suggests there to be two processes: one that dominates at the slow tempi (corresponding to the  $at^{-1}$  term) and another process that dominates at the faster tempi (the  $bt$  term). Further support for this two-term model comes from the observation that the former fit had an additional degree of freedom (three parameters) but accounted for less variance than the latter. This observation formed the basis for our model of timing deviations using two templates later on.

However, one could argue that the standard deviation of the intervals is not an appropriate measure, since the interval length itself varies across tempo. To compensate for this, we rescaled the unevenness to obtain the coefficient of variation (CV; calculated as standard deviation divided by mean interval length) (Fig. 1b). We found that instead of showing an optimum at medium tempo, the CV increased monotonically as a function of tempo, being in agreement with our first prediction. The CV increase was described by a quadratic model ( $R^2_{adj}=.81$ ,  $F(2,90)=182$ ,  $p<10^{-15}$ ) which was significantly superior to a linear model ( $F(1,82)=26.38$ ,  $p<10^{-5}$ ) (Fig. 1b). Again, the same was true for the other combinations of hand and playing direction (the quadratic model all  $F(2,90)>123.97$ ,  $p<10^{-20}$  and it

outperforms a linear fit, all  $F(1,82) > 6.20$ ,  $p < .02$ ). This means that playing becomes more uneven with increasing tempo, relative to the inter-keystroke duration.

In sum, the prediction of the generalised motor program hypothesis that unevenness should scale linearly with tempo, is shown to be false. A second prediction of this hypothesis is that the velocities of the keystrokes should become faster as tempo increases. This is because, by hypothesis, the same movements are executed in a smaller amount of time, resulting in a linear increase in the velocity with which the keys are depressed. We performed the same ANOVA as above with dependent variable keystroke velocity instead of temporal unevenness. We found a main effect of tempo ( $F(4,64) = 10.08$ ,  $p_{GG} < 10^{-5}$ ,  $\eta_G^2 = .18$ ) but not of hand ( $F(1,16) = .22$ ,  $p > .6$ ) nor direction ( $F(1,16) = 2.02$ ,  $p > .17$ ). There was an interaction between hand and direction ( $F(1,16) = 21.35$ ,  $p < .001$ ,  $\eta_G^2 = .001$ ) but no two-way interaction with tempo (both  $F(4,64) < 1.64$ ,  $p > .17$ ). The three-way interaction was significant ( $F(4,64) = 6.06$ ,  $p_{GG} < .004$ ,  $\eta_G^2 = .002$ ) and therefore we analysed the hand and direction combinations separately. For the right hand outward scales, the linear fit between tempo and velocity predicted by the generalised motor program hypothesis (shown in Fig. 1c) was indeed significant ( $F(1,83) = 17.29$ ,  $p < 10^{-4}$ ), but yielded only a marginal fit ( $R^2_{adj} = .16$ ). However, adding higher-order terms (starting with quadratic) did not yield a significantly better model ( $F(1,82) = 1.27$ ,  $p = .26$ ). In the other combinations of hand and direction, keystroke velocity always increased linearly with tempo (all  $F(1,83) > 13.02$ ,  $p < .001$ ) but adding a quadratic term did not improve fit (all  $F(1,82) < 1.30$ ,  $p > .25$ ). In sum, an approximately linear increase in keystroke velocity is apparent, as predicted by the generalised motor program hypothesis.

### **4.3.b Note-by-note irregularities: establishing the phrasal and neuromuscular templates**

Next, to answer the question what processes may be active across the different tempi we turn from the scale-level investigation to the note-by-note level (see methods). We calculated the note-by-note deviations and expressed these as a percentage of the interval length. We performed a Type-I repeated-measures ANOVA with 5 tempi x 2 directions x 2 hands x 15 notes as factors and irregularity as the dependent variable. There was a main

effect of note ( $F(14,224)=17.49$ ,  $p<10^{-28}$ ,  $\eta_G^2=.16$ ), reflecting that timing differed across notes. These differences will be the subject of our analysis in the remainder of this paper. There was no main effect of tempo ( $F(4,64)=.49$ ,  $p=.74$ ) nor of direction ( $F(1,16)=3.23$ ,  $p=.09$ ) or hand ( $F(1,16)=.19$ ,  $p=.66$ ). Hand and direction interacted ( $F(1,16)=4.93$ ,  $p=.04$ ) as did tempo and note ( $F(56,896)=8.05$ ,  $p<10^{-48}$ ,  $\eta_G^2=.11$ ) and hand and note ( $F(14,224)=3.30$ ,  $p<.001$ ,  $\eta_G^2=.01$ ). The former finding constitutes evidence against the generalised motor program hypothesis, which would have predicted no interaction of tempo and note. The other two-way interactions were not significant (both  $F<.71$ ,  $p>.58$ ). All the three-way interactions were significant (all  $F>3.72$ ,  $p<10^{-16}$ ) except for tempo, hand and body direction ( $F(4,64)=1.79$ ,  $p=.14$ ) and finally the four-way interaction between all factors was significant ( $F(56,896)=1.63$ ,  $p<.001$ ,  $\eta_G^2=.01$ ). As a result, we proceeded to analyse the tempo-note interactions for all four combinations of hand and direction separately.

First, we focus on the right-hand outward scales. We performed a two-way repeated-measures ANOVA with note and tempo as factors, which yielded a main effect of note ( $F(14,224)=7.95$ ,  $p<10^{-12}$ ,  $\eta_G^2=.17$ ), and an interaction of note and tempo ( $F(56,896)=9.73$ ,  $p<10^{-59}$ ), but no main effect of tempo ( $F(4,46)=.23$ ,  $p=.92$ ).

How does note timing vary across tempi? We found that at fast tempi, three distinct peaks in lateness appear (Fig. 2a, backward portion of the graph), corresponding to the positions in the two-octave scale where the index finger crosses the thumb (e.g., the index finger playing *g* after the thumb plays *f*). At slower tempi (Fig. 1b, front), an altogether different picture emerges, with the lateness taking the shape of an arc: the first few and last few notes are played late, whereas the rest are played slightly early. This observation, that essentially two qualitatively different templates are at play, forms the basis of the model that we provide below. That is, in the intermediate tempi, a gradual trade-off happens between these traces of the extreme tempi (Fig. 2a).

In order to arrive at a generative model of the timing deviations, in what follows we (a) established the timing profiles of the fast and slow tempi, and then (b) parameterised the trade-off between them across tempi.

Firstly we generated a simple model of the deviation trace at the slow tempo (2.7 notes/sec). We distinguished two phases (Fig. 3a): a phase where the deviation more or less

linearly decreases (notes  $c-g'$ ), i.e. speeding up in the middle, and then a phase where the deviation increases linearly, i.e., slowing down at the end ( $a'-c''$ ). That is, we fit a straight line to the trace from  $c$  to  $g'$ , and then a second line from  $a'$  to  $c''$ . This two-line model fitted the data well ( $R^2_{\text{adj}}=0.82$ ,  $\chi^2_{\text{red}}(12)=0.47$ ) (Fig. 3c). We took this fit quality based on only two parameters as an argument that there was no need to resort to more complex models of expressive timing to account for our data. In sum, our model of the slowest tempo implied a slow start of the scale playing, followed by a smooth speeding up and finishing with a dramatic slowing down. This latter shape is a widespread and consistent finding of musical phrase arcs in expressive playing (Friberg, Bresin, & Sundberg, 2006; Palmer & Krumhansl, 1987; Repp, 1990) and is taken to reveal that the two-octave scale is played as a single unit (Van Vugt et al., 2012) and we thus refer to this model as the *phrasal* template.

Next, we investigated the deviation trace of the fast tempo (13.3 notes/sec). Three distinct peaks appeared, corresponding to notes being played late (Fig. 3b). The peaks all corresponded to the notes played by the index finger, and, to a lesser degree, the preceding thumb keystrokes. Interestingly, the keystroke preceding each peak was held down less long, in a way that did not depend on tempo: Fig. 2b reveals that this pattern of key down times minus the interval length was more or less constant across tempi. This suggested that these timing deviations were dictated by the time the fingers needed to transition between the keystrokes (Engel et al., 1997; Furuya, Flanders, et al., 2011). It indicates that the deviations at this playing tempo were dominated by biomechanical constraints, and perhaps also by the hand anatomy (e.g. only the thumb can perform a three-dimensional rotation that optimises the horizontal translation of the hand position). Therefore we referred to this template as the *neuromuscular* template. To construct a model of this template, we started out with a deviation trace where all notes were on time. Then we introduced a delay  $d$  on the notes played by the index finger. To partially compensate for the introduced delay, the preceding note, played by the thumb, was also delayed by half that amount, that is  $d/2$ . Finally, the remaining notes appeared earlier so that the entire trace was centred around zero. In sum, the index finger was late, the preceding thumb was half as late and the other notes compensated for this lateness by being early. This model had a single parameter, but explained our data very well for so few degrees of freedom ( $R^2_{\text{adj}}=0.77$ ,  $\chi^2_{\text{red}}(12)=0.12$ ,  $p<.001$ ).

### 4.3.c Trade-off between the two templates

To investigate the extent to which we can approximate the temporal deviation traces of the intermediate tempi as a combination of the two extreme ones, we computed the least-squares linear combination of the two templates that best matched the deviation trace at that tempo. The overall fit ( $R^2_{\text{adj}}=0.69$ ,  $\chi^2_{\text{red}}(62)=0.43$ ,  $p<.0001$ ) of this model was good, given the few parameters used and the relatively crude fits that we used as templates. However, this fit did not completely capture the trade-off, since in this model an increase in one parameter did not necessarily entail a decrease in the other. Furthermore, it contained 10 degrees of freedom (2 parameters per tempo for 5 tempi). Instead, we decided to reduce the parameters to half by doing a single parameter fit per tempo. Given the deviation trace  $dt$  at a particular tempo, we computed an alternative least-square fit for each tempo, writing  $dt = \alpha * nm + (1-\alpha) * phr$ , where  $nm$  and  $phr$  were the vectors corresponding to our model fits of the neuromuscular and phrasal templates, respectively, and  $\alpha$  was the fitted parameter. The corresponding fit had only slightly reduced  $R^2_{\text{adj}}=.72$  ( $\chi^2_{\text{red}}(67)=0.43$ ,  $p<.0001$ ), even though we divided the amount of parameters by two (this latter fit is shown in Fig. 4). We take this as evidence for a trade-off between the two templates, rather than free variation between the two.

Then the question arose whether one could predict how this trade-off parameter ( $\alpha$ ) varies as a function of tempo. To this effect, we fit a quartic function to specify  $\alpha$  as a function of tempo as follows:  $\alpha(t)=at^4$ , where  $t$  is the tempo (in strokes/sec). We fitted for the parameter  $a$ . The choice for a quartic function rather than any other monotonic function is essentially arbitrary. With this new model, we reduced the parameters to one ( $a$ ) instead of 4, at minimal loss in explained variance ( $R^2_{\text{adj}}=0.70$ ,  $\chi^2_{\text{red}}(70)=0.40$ ,  $p<10^{-7}$ ). If we fit only the neuromuscular template, the obtained model has a poorer fit ( $R^2_{\text{adj}}=0.40$ ,  $\chi^2_{\text{red}}(72)=0.76$ ,  $p>.06$ ), and the same is true for the phrasal template ( $R^2_{\text{adj}}=0.18$ ,  $\chi^2_{\text{red}}(71)=0.55$ ,  $p<.001$ ). This latter result indicates that the two templates both contribute to the pattern of timing deviations observed in these pianists.

In sum, we present a generative model that predicts the temporal deviation trace in scale playing based on two templates that we label as phrasal and neuromuscular and a trade-off function using a single-parameter ( $a$ ). The phrasal template is fit using two parameters (the slopes of the two lines), and the neuromuscular template is based on a

single parameter (the lateness of the index finger passage). Using only these four parameters, we predict the deviation traces of 5 tempi times 15 notes, that is, 75 data points. We have obtained an  $R^2_{\text{adj}}$  that shows we cover not all variability, but still a significant portion of it ( $R^2_{\text{adj}}=0.70$ ).

#### 4.3.d Generalising to all outward scales

So far, we have generated a model for the timing deviations of the right hand outward scales. However, we have understood little of how general this model is. As a cross-validation of our model, we now turn to the left hand outward scales (Fig. 5 right hand traces). When the left hand plays descending scales, the sequence of movements is exactly the same as the right hand ascending scales we have analysed thus far (outward), except in that they are mirrored in space. If the two templates established so far are effector-unspecific representations, the prediction is therefore that the model will produce a fit to the left hand outward scales when mirrored in space. Indeed, if we apply the very same fit as we have obtained for the right hand scales to the left hand outward scales, we obtain a significant model ( $R^2_{\text{adj}}=0.39$ ,  $\chi^2_{\text{red}}(70)=0.67$ ,  $p=.01$ ). We emphasise at this point, that the parameters in this model have been fit for a separate subset of our data (i.e. the right hand outward scales). We feel that the fact that this model even with its parameters intact generalises, supports our analysis. Furthermore, if we take the outward movements together, our model has an overall fit ( $R^2_{\text{adj}}=0.55$ ,  $\chi^2_{\text{red}}(145)=0.52$ ,  $p<10^{-7}$ ). The model could be further optimised by fitting for the two hands separately.

#### 4.3.e Generalising to inward scales

Now inward scales are a different case. Those who have played scales on a piano know that performing the same (inward or outward) movement with both hands at the same time is easy, whereas performing different movements (one hand inward, the other outward) is much harder. Let us first turn to the fastest tempo, at which the neuromuscular template dominates. To understand what temporal deviation trace is predicted for the inward movement, imagine the following. If we capture the outward scale on video and

play it backwards, we would witness the inward movement. How would the temporal deviation trace for the original and the backwards variant relate? The reader is invited to convince himself that, first of all, the corresponding deviations would be reversed in time. Next, they would be inverted in lateness: notes that were late in the one movement would be early in the time-mirrored version, and vice versa. Finally, one further operation is required to make the deviation traces comparable. We argued previously that the timing pattern found at the fastest tempo is determined by the difficulty of the thumb- and index-finger transitions. That is, the keystroke after this transition will be affected (i.e. late) because of the difficulty of the transition. This means that when the order of the keystrokes is reversed, the keystroke affected by the transition delay is one keystroke earlier. In other words, we expect the timing pattern to be shifted by one note forward in its entirety. We discard the last note of the inward trace (since there is no corresponding note in the outward trace). Since our traces are centered around zero due to the line fit that obtained them, the timing of the new data point at the beginning of the trace is fixed (and it is the same as the discarded note at the end of the trace).

In sum, we take the deviation trace of the right hand outward movement at the fastest tempo, reverse the order of the notes, invert their lateness (i.e. multiply the deviations by -1), and then shift the entire trace by one note, the corresponding trace correlates highly with the right hand inward scales (Spearman  $\rho(14)=.75$ ,  $p=.001$ ). Similarly, for the left hand scales (Spearman  $\rho(14)=.73$ ,  $p=.001$ ). Figure 6 illustrates the surprising match between the traces after this manipulation (compare with Fig. 5 bottom left and right).

Next, we turn to the slowest tempo, at which we previously showed the phrasal template dominates. In this case, inverting the trace in time as shown above does not yield a good fit for the left hand ( $\rho(14)=-.34$ ,  $p=.89$ ) nor for the right hand ( $\rho(14)=-.63$ ,  $p=.99$ ). On the contrary, the non-inverted, original (time-invariant) traces correlates highly for the left ( $\rho(14)=.75$ ,  $p<.001$ ) and right hand ( $\rho(14)=.81$ ,  $p=.0001$ ). In other words, although at the fast tempo the timing deviations are mirrored in time (Fig. 5 bottom plots), at the slower tempo the timing deviations are invariant to time inversion (Fig. 5 top plots). As a result, the intermediate tempi, which are a linear combination of these two motor primitives, at face value reveal no obvious relation (Fig. 5 middle plots). Only our decomposition of the timing

deviations in two motor primitives has revealed the underlying symmetries in time and space. Our study therefore yields the lesson that symmetries in movements may be readily revealed when movements are decomposed into motor primitives.

In order to model the inward movements across all tempi, the predictions are the following. The neuromuscular template is inverted in time, whereas the phrasal template remains the invariant. Taking the same parameter settings as for the right hand outward scales we were able to generate the expected timing pattern, which fit the totality of the inward dataset well ( $R^2_{\text{adj}}=0.42$ ,  $\chi^2_{\text{red}}(149)=.55$ ,  $p<.001$ ). The explained variance is not overwhelming, but notice that we have generalised from one movement to its mirror image in time, two movements that on the surface appear completely unrelated, and our parameters were fit on a different subset of the data (the right hand outward scales).

#### **4.3.f Generalising to both hands and directions**

Finally, consider all data points in our data set: the four combinations of hand and direction. Using our model that was fit exclusively on the right hand outward scale, as a cross-validation we generalised to the other directions and hands as above. The overall model contains four parameters but predicts temporal deviations for 15 notes across 5 tempi, 2 hands and 2 directions, that is, 300 data points. The model is highly significant ( $R^2_{\text{adj}}=0.49$ ,  $\chi^2_{\text{red}}(299)=0.53$ ,  $p<10^{-7}$ ) and its fit could potentially be further optimised by fitting the parameters on the entire dataset, instead of only on the subset of right hand outward scales. However, we find the current analysis more convincing, since the remainder of the dataset serves as cross-validation.

#### **4.3.g Individual level and correlation with accumulated playing time**

Our analysis so far has investigated the mean deviation trace of the pianists taken together and has revealed a trade-off between two movement templates across tempi. However, the question that remains is to what extent this trade-off is present in individual pianists and whether it depends on how experienced the pianist is.

Therefore, in the final phase of our analysis, we performed our model fit for each

pianist individually. Given a pianist and a tempo, we computed the mean deviation trace ( $dt$ ) as before, and we approximate it as a linear combination of the neuromuscular ( $nm$ ) and phrasal template ( $phr$ ) as follows:  $dt = nm(d) + phr(i, s_1, s_2)$ . Here the neuromuscular model  $nm$  was defined by one parameter ( $d$ ) and it was constructed precisely as specified before. Furthermore, the phrasal model was again defined as two lines, one descending line for notes 1:12 ( $c-g$ ) and one ascending line for notes 13:15 ( $a'-c''$ ). That is, we have three parameters: the intercept ( $i$ ) and slope ( $s_1$ ) of the first line, and the slope of the second line ( $s_2$ ) (note that the intercept of the second line is fixed because this line is forced to intersect with the first one at  $a'$ ).

That is, we have four parameters for each tempo in the trace, which we estimate through a non-linear least-square fit through the Nelder-Mead simplex method. Additionally, we forced the values of  $s_1$  to be negative and  $d$  and  $s_2$  to be positive by introducing corresponding penalties in the squared error term of the optimisation. We then calculated the amount of variance (in the deviation trace) that is explained by each of the two templates by subtracting the summed squared error (SSE) of the two-template model from the SSE of the other (single) template model. Note that the difference between this analysis and the trade-off model (introduced in the previous part of this paper) is that the latter is based on the average traces (and their SD) of all participants whereas the former is fit for each participant individually.

The result corroborates our previous finding (Fig. 7): the phrasal template explains most variance in the slower tempi, the neuromuscular template most of the variance in the higher tempi. At the intermediate tempi there is a clear trade-off.

Finally, we turn to the question how this amount of variance explained by the two tempi relates to the total accumulated playing time of the pianist (in hours). For each pianist, we summed the amount of variance explained by the neuromuscular template across the different tempi (in percentage of the total variance). We hypothesised that for more experienced pianists, the contribution of the neuromuscular template would decrease. Surprisingly, the opposite correlation was found (Fig. 8): more experienced pianists' playing showed a greater contribution of the neuromuscular template (Spearman  $\rho(14)=.56$ ,  $p=0.03$ ). Even eliminating the one pianist with over 36 thousand hours, the correlation was significant (Spearman  $\rho(13)=.59$ ,  $p=.03$ ). There was no correlation between the accumulated

practice hours of the pianist and the variance explained by the phrasal template (Spearman  $\rho(14)=-.30$ ,  $p=.28$ ). On the other hand, the amount of discarded note material (in %) did not correlate with the total amount of variance explained by the neuromuscular template (in %) (Spearman  $\rho(14)=-.28$ ,  $p=.31$ ). However, it did correlate significantly with the contribution of the phrasal template (Spearman  $\rho(14)=.64$ ,  $p=.01$ ). There was no correlation between the accumulated practice hours and the amount of residual variance, not explained by either of the two templates (Spearman  $\rho(14)=-.21$ ,  $p=.46$ ).

#### 4.4 Discussion

The present study aimed to clarify the motor control processes that govern motor productions at various speeds. We first of all replicated the finding that scale playing is most even at around 8 keystrokes/sec, i.e. playing is most precise at that tempo. This was previously interpreted as showing that too slow or too fast tempi made it difficult to establish a playing rhythm (Wagner, 1971). However, our investigation of the coefficient of variation suggests that when one takes into account the difference in speed between the conditions, unevenness monotonically increases with tempo. Taken together with the finding that the keystroke velocities increased, this suggests that the variation is partly due to signal dependent noise (SDN) included in motor commands issued to muscles (Faisal, Selen, & Wolpert, 2008). SDN has been shown to increase as a function of muscular force (Jones, Hamilton, & Wolpert, 2002), in line with the speed-accuracy trade-off (Fitts, 1954).

How does the motor system implement the same movement at various speeds? Previous studies have advanced the hypothesis of generalised motor programs (Shapiro et al., 1981), predicting that the timing of motor output scales with the tempo at which the motor program is executed, i.e., relational invariance (Repp, 1994). Our findings speak against such a construction in the case of piano scale playing. Timing deviations were qualitatively different at fast and slow tempi, consistent with previous findings (MacKenzie & Van Eerd, 1990). Indeed, an interaction effect of note and tempo was found.

The unevenness of scale playing (defined as the standard deviation of the

inter-keystroke intervals) has been defined on the level of the scale as a whole (Jabusch et al., 2004). However, calculating and using this quantity tacitly assumes that the temporal deviations of the individual notes are independent and identically distributed. Our data note-level analysis revealed that these temporal deviations are systematic and show qualitatively different patterns as a function of tempo. These temporal deviations can be quite well explained as a combination of two qualitatively different templates. We have shown that taking only one of these templates and parametrically modulating it as a function of tempo was not enough to account for the data. This observation constitutes evidence against the relational invariance prediction derived from the generalised motor program hypothesis.

We now take time to look deeper into the motor primitives that we identified. During typical playing, the thumb passes underneath the index and middle finger at three distinct points during the two-octave scale (thumb-passage) and then the index finger has to pass over the thumb (index-passage). At fast tempi, a template that instantiates the delays inherent in this so-called thumb-under manoeuvre (Engel et al., 1997) is dominant. As a result, we labelled this template as the neuromuscular template. In conjunction with the finding that key down times preceding these delayed keystrokes are systematically shorter, this template can be roughly explained from low-level constraints on finger movements. A fixed amount of time is required for the fingers to translate to their next keystroke position. Interestingly, this translation time is independent of distance: it is the same when the thumb underneath two fingers (between  $e$  and  $f$  and  $e'$  and  $f'$ ) or three fingers (between  $b$  and  $c'$ ) (Fig. 3b). Interestingly, there appears to be a tendency for the transition time to be actually *less* for the longer movement in the middle (Fig. 6). We do not have an explanation for this finding. The resulting temporal deviations show up convincingly at the fast tempi (Fig. 2) and they are modelled well by a single-parameter model introduced above.

At slower tempi, however, the pattern of deviations shows two clear landmarks: mildly slow playing in the beginning and a more pronounced slowing at the end. Timing traces of this shape are a widespread finding in musical timing literature (Palmer & Krumhansl, 1987; van Vugt et al., 2012) and appears in speech timing for languages across the globe (Turk & Shattuck-Hufnagel, 2007). It is similarly a typical feature of locomotion to

slow down at the end in a very similar fashion (Friberg & Sundberg, 1999). That is, we interpret this template as an expressive device that signals phrasal structure. Its dominance at slower tempi reveals that since the temporal constraints on playing are less strict, pianists may shift their cognitive resources to generating an aesthetically pleasing sound (Friberg & Sundberg, 1999). Furthermore, the fact that this movement template is absent at the fastest tempo suggests that it reflects not a fundamental feature movement but rather a control property that is present whenever temporal constraints in the context allow it. Interestingly, a similar phenomenon appears to hold in action observation, where continuing motion appears to slow down as a result of our perceptual system's adaptation to it (Goldstein, 1957) and in general a salient event (such as a start or end of a scale sequence in our case) slows down our perception of time (Tse, Intriligator, Rivest, & Cavanagh, 2004). However, we argue that this action observation phenomenon cannot account for the timing deviations observed in our phrasal template. Pianists were instructed to aim for temporal evenness and therefore if the beginning and end of the scale sounded slower, they would respond by speeding up in compensation.

It is an important finding that two templates manage to account for a wide variety of tempi. Our five-parameter model has managed to cover a large portion of the variance in this complex dataset. We feel this reflects not so much a feat of our analysis method, but of the motor system employing a dimensionality reduction of control (Bernshtein, 1967). The motor system is capable of generating rich repertoires of movements to fulfil various task demands. This flexibility, however, causes neural processing to become intractable when having to produce fast sequential movements. A simplification of movement control through combining a small number of fundamental motor patterns (i.e. primitives) has been evident in various simple and complex motor behaviours, such as kicking (D' Avella & Bizzi, 2005; Hart & Giszter, 2004), walking (Ivanenko, Poppele, & Lacquaniti, 2004), reaching (D' Avella et al., 2006), hand grasping (Gentner & Classen, 2006; Santello, Flanders, & Soechting, 2002) and musical performance (Furuya, Flanders, et al., 2011; Gentner et al., 2010). Together with the current finding, it seems likely that the nervous system tends to efficiently produce a wider repertoire of movements using a smaller set of motor primitives.

Our study then reveals the motor system's capacity to combine the two templates at intermediate tempi and manipulate them through spatial and temporal inversions. Contrary

to the dominant understanding of generalised motor programs we show that the motor primitives under consideration here show a remarkable flexibility in that they can be mirrored in space as well as in time. This result suggests that the motor system may first increase control complexity by using two instead of one motor primitives (contrary to generalised motor program accounts) but is then able to flexibly adapt these primitives to different task constraints by mirroring them in space and time. This latter flexibility has, to our knowledge, not been proposed previously. Furthermore, the linear combination of a time-inverted and time-invariant motor primitive at intermediate tempi results in a motor pattern that is symmetric in neither dimension, revealing the complexity of control that can be achieved through combination and mirroring of simple primitives.

Future studies could further clarify the process by which the motor system is able to recombine motor primitives through spatial and time inversion. For example, to what extent would motor learning in participants trained on a movement sequence be able to generalise to performing the same movement mirrored in space or in time?

In sum, modular control more or less generalised across the hands and across the movement directions. A similarity in both phrasal and neuromuscular templates across the hands may further simplify the control of sequential movements. Our results indicate a great degree of overlap between motor representations used to control the both hands (Fig. 5), although, particularly at faster tempi, the transition-induced hesitations due to the thumb-under manoeuvre appear increased in the left hand. Indeed, previous studies revealed inter-manual differences due to inter-segmental dynamics between the dominant and non-dominant hands (Bagesteiro & Sainburg, 2002; Heuer, 2007; Sainburg, 2002; Sainburg & Kalakanis, 2000). Indeed, during the thumb-under manoeuvre the proximal body portion such as the wrist moves predictively to ensure smooth translation of the hand (Engel et al., 1997). This motion possibly creates the inter-segmental dynamics at the hand, which originates from mechanical interaction across musculatures and matter. The nervous system has to compensate for this perturbing dynamics in order to accomplish the thumb-under successfully. A better compensation of the inter-segmental dynamics by the dominant hand compared with the non-dominant one (Sainburg, 2005) may underlie our finding of less pronounced thumb-under peaks in the dominant hand (Fig. 6).

Production of a planned movement necessitates compensation for neuromuscular

constraints on the limb (Kawato, 1999). This requires accurate representation of the limb dynamics, referred to as an internal inverse model, which is acquired through extensive training (Osu et al., 2002; Thoroughman & Shadmehr, 2000). However, our finding of the timing deviations captured by the neuromuscular template of skilled pianists indicated failure to compensate for biomechanical constraints of the hand particularly when playing at fast tempo. Indeed, on the individual level, surprisingly, more advanced pianists showed a greater contribution of neuromuscular constraints in their timing deviations. It is thus unlikely that years of piano training can form an internal model that accounts for the hand dynamics changing in relation to various contexts of movements, such as forces originating from inertia and viscosity of the hand during the thumb-under. One possible explanation is that more advanced pianists managed to reduce sources of variance other than those due to the neuromuscular template. In addition, the average timing deviation of 10% is just below the detection threshold of deviations in an isochronous sequence: approximately 15% at a 80ms IOI sequence (Ehrlé & Samson, 2005). Indeed, it amounted to approximately 8 msec at the fastest tempo, which is below the threshold of detectability in scale timing differences by experienced listeners (Van Vugt et al., in revision). This is furthermore in line with the 10 msec threshold for detecting temporal asynchronies between the onset of two notes (Mossbridge, Fitzgerald, O'Connor, & Wright, 2006) as well as for detecting deviations from a tempo (Repp, 1992). It is thus possible that the present skilled pianists took maximal advantage of this perceptual allowance to cope with the biomechanical constraints that emerge during the thumb-under, which can be acquired through representing both perceptual threshold and hand dynamics. Similarly, the correlation between the phrasal template contribution and the amount of errors suggests that the phrasal template may be a mark of less experienced pianists (as evidenced by amount of wrong notes). However, a limitation of the current study is that our pool of participants consisted exclusively of highly skilled pianists with more than 10,000 practice hours. Future research could apply the analysis introduced here to less skilled pianists to gain insight in the role of the phrasal and neuromuscular timing deviations in the early stages of musical practice.

Our results also open the road to a deeper investigation of anomalies in scale playing (or perhaps other musical materials) in patients with musician's dystonia (MD). MD is characterised by structural and functional maladaptation at cortical and subcortical

regions, which result in a loss of fine motor control (Altenmüller & Jabusch, 2009). Scale playing unevenness has been shown a meaningful indicator of MD severity (Jabusch et al., 2004; Rosenkranz, Butler, Williamon, & Rothwell, 2009) and correlated to neuro-anatomical irregularities (Granert, Peller, Jabusch, Altenmüller, & Siebner, 2011). Based on the previous findings we can now ask what qualitative aspects of playing are altered in MD patients. Furthermore, are the motor primitives themselves affected or is it their recombination of them that is altered in these patients? The present study provides a clue how these issues may be tackled straightforwardly. Furthermore, the current study provides a basis for identifying symptoms specific to other movement disorders such as tendonitis and carpal tunnel syndrome, which have been prevalent among pianists (Hochberg, Leffert, Heller, & Merriman, 1983).

## 4.5 References

- Altenmüller, E., & Jabusch, H.-C. (2009). Focal hand dystonia in musicians: phenomenology, etiology, and psychological trigger factors. *Journal of Hand Therapy, 22*(2), 144–154. doi:10.1016/j.jht.2008.11.007
- Angelaki, D. E., & Soechting, J. F. (1993). Non-uniform temporal scaling of hand and finger kinematics during typing. *Experimental Brain Research, 95*(2), 319–329.
- Bagesteiro, L. B., & Sainburg, R. L. (2002). Handedness: dominant arm advantages in control of limb dynamics. *Journal of neurophysiology, 88*(5), 2408–2421. doi:10.1152/jn.00901.2001
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods, 37*(3), 379–384.
- Bernshtein, N. A. (1967). *The co-ordination and regulation of movements*. Pergamon Press.
- Bhatara, A., Tirovolas, A. K., Duan, L. M., Levy, B., & Levitin, D. J. (2011). Perception of emotional expression in musical performance. *Journal of Experimental Psychology. Human Perception and Performance, 921–934*. doi:10.1037/a0021922
- Clarke, E. F. (1982). Timing in the performance of Erik Satie’s “Vexations.” *Acta Psychologica, 50*(1), 1–19. doi:10.1016/0001-6918(82)90047-6
- Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition, 80*(3), B1–10.
- D’ Avella, A., & Bizzi, E. (2005). Shared and specific muscle synergies in natural motor behaviors. *Proceedings of the National Academy of Sciences of the United States of America, 102*(8), 3076–3081. doi:10.1073/pnas.0500199102
- D’ Avella, A., Portone, A., Fernandez, L., & Lacquaniti, F. (2006). Control of fast-reaching movements by muscle synergy combinations. *The Journal of Neuroscience, 26*(30), 7791–7810. doi:10.1523/jneurosci.0830-06.2006
- Desain, P., & Honing, H. (1994). Does expressive timing in music performance scale proportionally with tempo? *Psychological Research, 56*(4), 285–292. doi:10.1007/BF00419658
- Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: influence of the tempo and musical backgrounds of listeners. *Brain and Cognition, 58*(1), 133–147. doi:10.1016/j.bandc.2004.09.014
- Engel, K. C., Flanders, M., & Soechting, J. F. (1997). Anticipatory and sequential motor control in piano playing. *Experimental Brain Research, 113*(2), 189–199.
- Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9*(4), 292–303. doi:10.1038/nrn2258
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381–391. doi:10.1037/h0055392
- Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology, 2*(2), 145–161.
- Friberg, A., & Sundberg, J. (1999). Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners. *The Journal of the Acoustical Society of America, 105*(3), 1469–1484. doi:10.1121/1.426687
- Furuya, S., Aoki, T., Nakahara, H., & Kinoshita, H. (2012). Individual differences in the biomechanical effect of loudness and tempo on upper-limb movements during repetitive piano keystrokes. *Human Movement Science, 31*(1), 26–39. doi:10.1016/j.humov.2011.01.002
- Furuya, S., Flanders, M., & Soechting, J. F. (2011). Hand kinematics of piano playing. *Journal of Neurophysiology, 106*(6), 2849–2864. doi:10.1152/jn.00378.2011
- Furuya, S., Goda, T., Katayose, H., Miwa, H., & Nagata, N. (2011). Distinct inter-joint coordination during fast

- alternate keystrokes in pianists with superior skill. *Frontiers in Human Neuroscience*, 5, 50. doi:10.3389/fnhum.2011.00050
- Furuya, S., & Soechting, J. F. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. *Experimental Brain Research*, 204(2), 223–237. doi:10.1007/s00221-010-2307-2
- Gentner, R., & Classen, J. (2006). Modular Organization of Finger Movements by the Human Central Nervous System. *Neuron*, 52(4), 731–742. doi:10.1016/j.neuron.2006.09.038
- Gentner, R., Gorges, S., Weise, D., aufm Kampe, K., Buttman, M., & Classen, J. (2010). Encoding of motor skill in the corticomuscular system of musicians. *Current Biology*, 20(20), 1869–1874. doi:10.1016/j.cub.2010.09.045
- Goldstein, A. G. (1957). Judgments of visual velocity as a function of length of observation time. *Journal of Experimental Psychology*, 54(6), 457–461.
- Granert, O., Peller, M., Jabusch, H.-C., Altenmüller, E., & Siebner, H. R. (2011). Sensorimotor skills and focal dystonia are linked to putaminal grey-matter volume in pianists. *Journal of Neurology, Neurosurgery & Psychiatry*, 1225–1231. doi:10.1136/jnnp.2011.245811
- Hart, C. B., & Giszter, S. F. (2004). Modular premotor drives and unit bursts as primitives for frog motor behaviors. *The Journal of Neuroscience*, 24(22), 5269–5282. doi:10.1523/jneurosci.5626-03.2004
- Heuer, H. (2007). Control of the dominant and nondominant hand: exploitation and taming of nonmuscular forces. *Experimental brain research*, 178(3), 363–373. doi:10.1007/s00221-006-0747-5
- Hochberg, F. H., Leffert, R. D., Heller, M. D., & Merriman, L. (1983). Hand difficulties among musicians. *JAMA: the journal of the American Medical Association*, 249(14), 1869–1872.
- Hollerbach, M. J., & Flash, T. (1982). Dynamic interactions between limb segments during planar arm movement. *Biological Cybernetics*, 44(1), 67–77.
- Hore, J., O'Brien, M., & Watts, S. (2005). Control of joint rotations in overarm throws of different speeds made by dominant and nondominant arms. *Journal of Neurophysiology*, 94(6), 3975–3986. doi:10.1152/jn.00327.2005
- Ivanenko, Y. P., Poppele, R. E., & Lacquaniti, F. (2004). Five basic muscle activation patterns account for muscle activity during human locomotion. *The Journal of physiology*, 556(Pt 1), 267–282. doi:10.1113/jphysiol.2003.057174
- Jabusch, H.-C., Vauth, H., & Altenmüller, E. (2004). Quantification of focal dystonia in pianists using scale analysis. *Movement Disorders*, 19(2), 171–180. doi:10.1002/mds.10671
- Jones, K. E., Hamilton, A. F., & Wolpert, D. M. (2002). Sources of signal-dependent noise during isometric force production. *Journal of Neurophysiology*, 88(3), 1533–1544.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6), 718–727.
- Khalfa, S., Roy, M., Rainville, P., Dalla Bella, S., & Peretz, I. (2008). Role of tempo entrainment in psychophysiological differentiation of happy and sad music? *International Journal of Psychophysiology*, 68(1), 17–26. doi:10.1016/j.ijpsycho.2007.12.001
- Kopiez, R., Jabusch, H.-C., Galley, N., Homann, J.-C., Lehmann, A. C., & Altenmüller, E. (2011). No disadvantage for left-handed musicians: The relationship between handedness, perceived constraints and performance-related skills in string players and pianists. *Psychology of Music*, 40(3), 357–384. doi:10.1177/0305735610394708
- Lacquaniti, F., Ferrigno, G., Pedotti, A., Soechting, J. F., & Terzuolo, C. (1987). Changes in spatial scale in drawing and handwriting: kinematic contributions by proximal and distal joints. *The Journal of Neuroscience*, 7(3), 819–828.
- MacKenzie, C. L., & Van Eerd, D. L. (1990). Rhythmic precision in the performance of piano scales: Motor psychophysics and motor programming. *Attention and Performance*, 13(Hillsdale, NJ), 375–408.
- Mossbridge, J. A., Fitzgerald, M. B., O'Connor, E. S., & Wright, B. A. (2006). Perceptual-learning evidence for separate processing of asynchrony and order tasks. *Journal of Neuroscience*, 26(49), 12708–12716. doi:10.1523/jneurosci.2254-06.2006
- Osu, R., Franklin, D. W., Kato, H., Gomi, H., Domen, K., Yoshioka, T., & Kawato, M. (2002). Short- and

- long-term changes in joint co-contraction associated with motor learning as revealed from surface EMG. *Journal of Neurophysiology*, 88(2), 991–1004.
- Palmer, C. (1997). Music Performance. *Annual Review of Psychology*, 48, 115–138. doi:10.1146/annurev.psych.48.1.115
- Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 116–126. doi:10.1037/0096-1523.13.1.116
- Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *The Journal of the Acoustical Society of America*, 88(2), 622–641.
- Repp, B. H. (1992). Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Cognition*, 44(3), 241–281.
- Repp, B. H. (1994). Relational invariance of expressive microstructure across global tempo changes in music performance: an exploratory study. *Psychological Research*, 56(4), 269–284.
- Rosenkranz, K., Butler, K., Williamson, A., & Rothwell, J. C. (2009). Regaining motor control in musician's dystonia by restoring sensorimotor organization. *Journal of Neuroscience*, 29(46), 14627–14636. doi:10.1523/jneurosci.2094-09.2009
- Sainburg, R. L. (2002). Evidence for a dynamic-dominance hypothesis of handedness. *Experimental Brain Research*, 142(2), 241–258. doi:10.1007/s00221-001-0913-8
- Sainburg, R. L. (2005). Handedness: differential specializations for control of trajectory and position. *Exercise and sport sciences reviews*, 33(4), 206–213.
- Sainburg, R. L., & Kalakanis, D. (2000). Differences in control of limb dynamics during dominant and nondominant arm reaching. *Journal of Neurophysiology*, 83(5), 2661–2675.
- Santello, M., Flanders, M., & Soechting, J. F. (2002). Patterns of hand motion during grasping and the influence of sensory guidance. *The Journal of Neuroscience*, 22(4), 1426–1435.
- Shapiro, D. C., Zernicke, R. F., & Gregor, R. J. (1981). Evidence for generalized motor programs using gait pattern analysis. *Journal of Motor Behavior*, 13(1), 33–47.
- Summers, J. J. (1975). The role of timing in motor program representation. *Journal of Motor Behavior*, 7, 229–241.
- Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, 407(6805), 742–747. doi:10.1038/35037588
- Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Attention, Perception, & Psychophysics*, 66(7), 1171–1189.
- Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, 35(4), 445–472. doi:10.1016/j.wocn.2006.12.001
- Van Vugt, F. T., Jabusch, H.-C., & Altenmüller, E. (2012). Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Frontiers in Auditory Cognitive Neuroscience*, 3, 495. doi:10.3389/fpsyg.2012.00495
- Van Vugt, F. T., Jabusch, H.-C., & Altenmüller, E. (2013). Individuality that is unheard of: systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint. *Frontiers in Cognitive Science*, (4), 134. doi:10.3389/fpsyg.2013.00134
- Wagner, C. (1971). The influence of the tempo of playing on the rhythmic structure studied at pianist's playing scales. In J. Vredenburg & J. Wartenweiler (Eds.), *Medicine and sport* (Vol. 6, pp. 129–132). Basel: Karger.

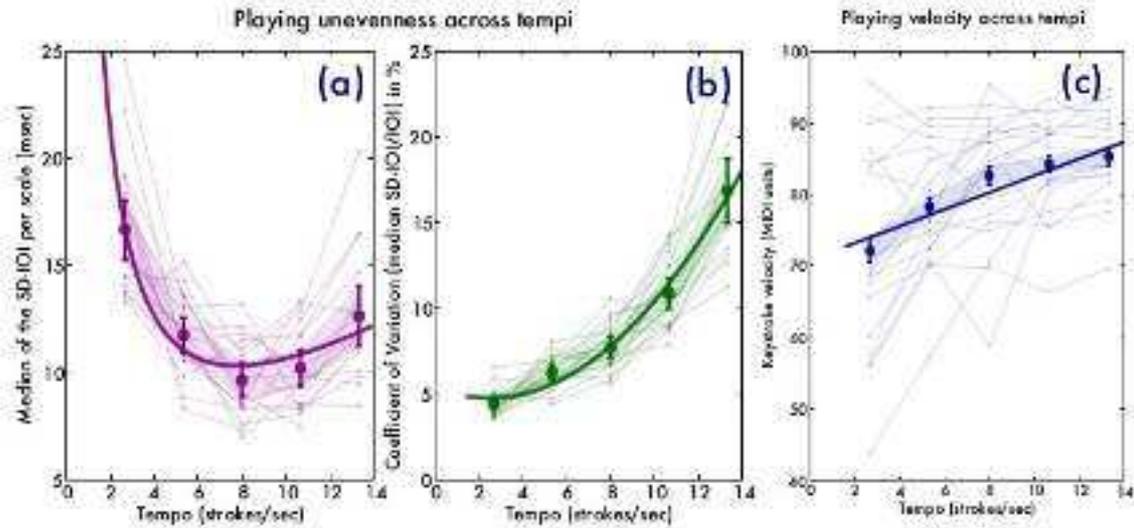


Figure 1. (a) Playing unevenness (standard deviation of the inter-keystroke-interval) as a function of tempo for the right-hand outward scales. The faded lines indicate the individual pianists and the error bars represent the 2\*standard error of the mean across all subjects. The solid line represents our two-process model fit. (b) Rescaled representation of the unevenness (coefficient of variation calculated as standard deviation divided by the mean inter-keystroke-interval). Error bars indicate 2\*standard error of the mean. The solid line represents our quadratic model. (c) Mean keystroke velocity (resulting in loudness) across tempi. The solid line indicates a linear model fit.

Sounds on time - Over-learning in Musical Experts

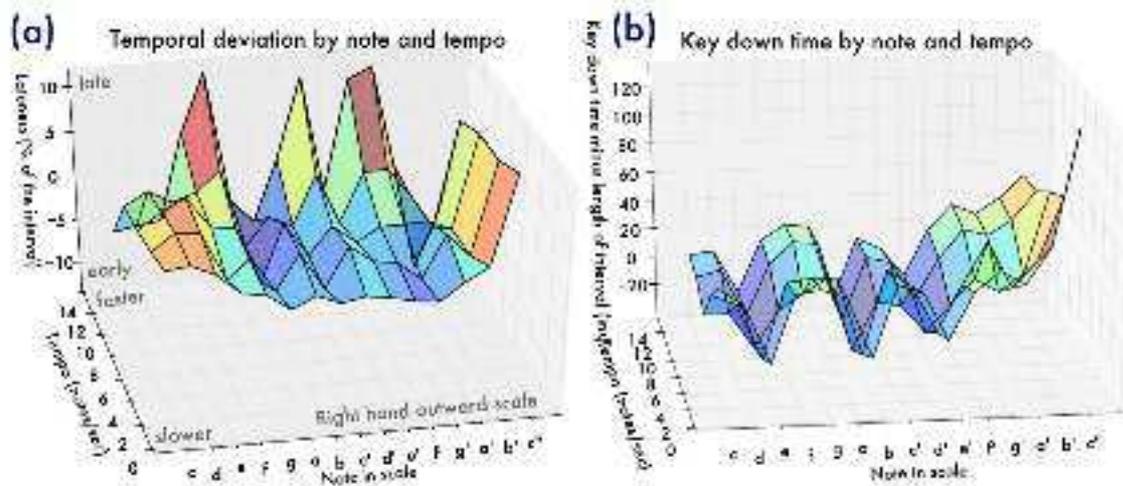


Figure 2. (a) The mean deviation (lateness) in percentage of the interval (right hand outward scales). The depth-axis shows the different tempi. Colour coding is added that reflects the mean vertical position of each graph segment. The deviation pattern is qualitatively different for the slow and fast tempi. (b) the key down time minus the target inter-keystroke interval (in msec), revealing that the keys preceding a thumb-passage are released earlier by an amount of time that is more or less constant across tempi.

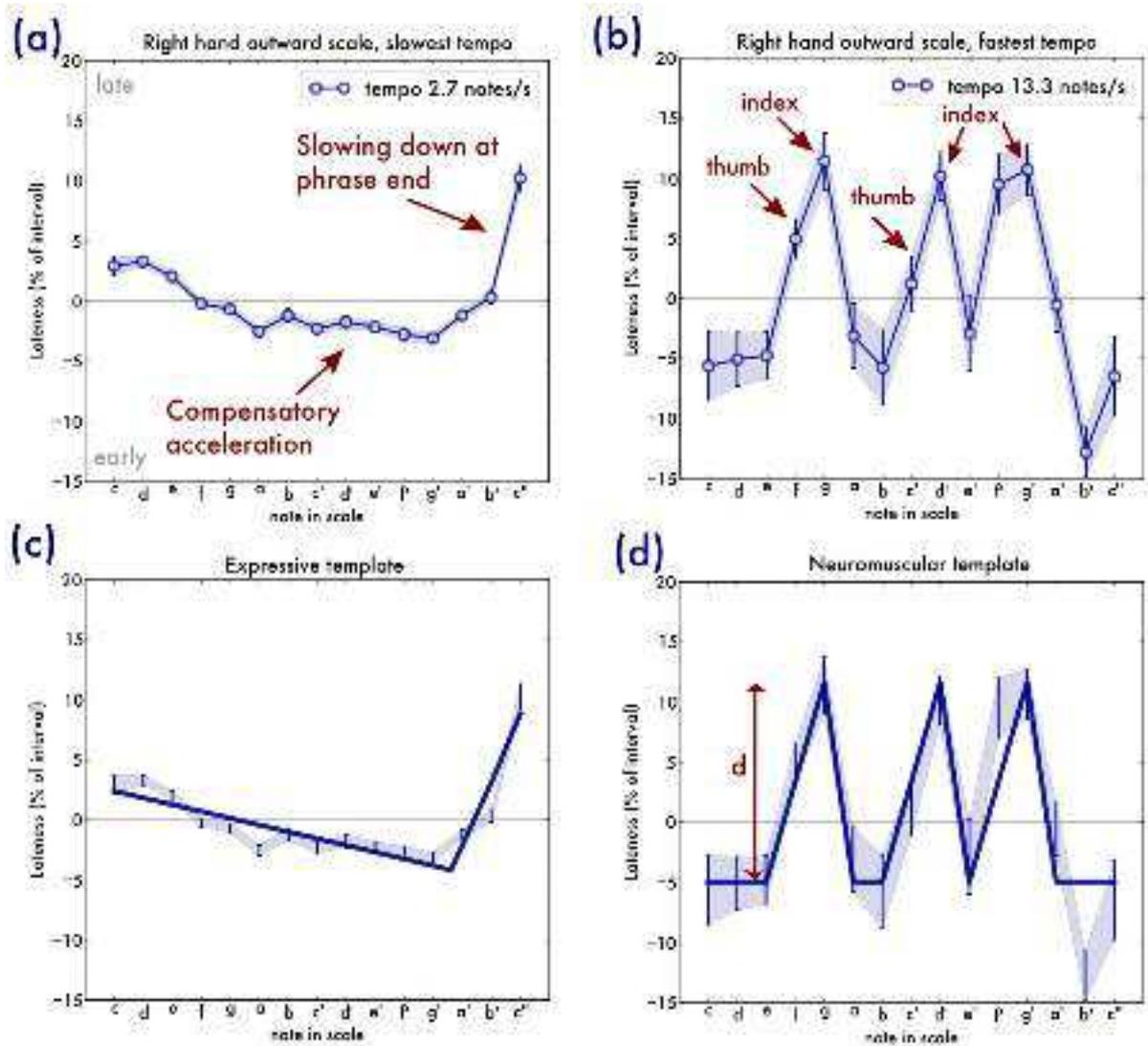


Figure 3. (a) Mean lateness (expressed as coefficient of variation in %) for the slowest playing tempo. (b) Mean lateness for the fastest tempo. Error bars and shaded area indicate standard error across subjects. (c) Our fit to the deviation trace of the slow tempo. The solid line shows our model. (d) Our fit at the fastest tempo. The error bars represent the standard error of the mean.

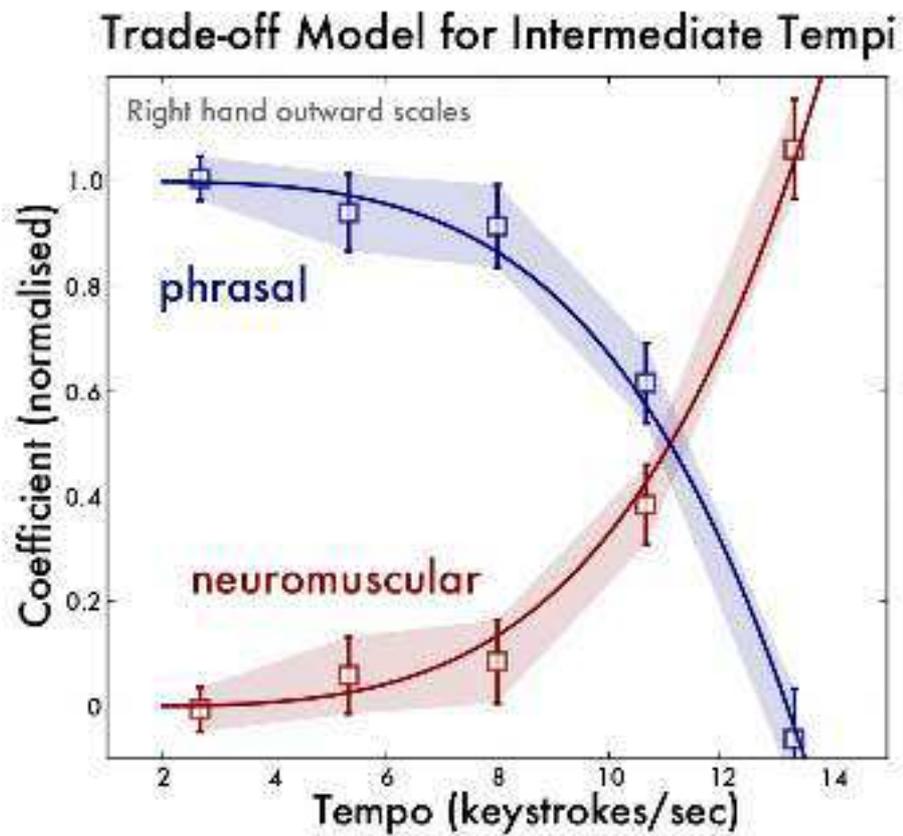


Figure 4. The squares and their error bars indicate the result of a least-square fit  $\alpha(t) \cdot nm + (1 - \alpha(t)) \cdot phr$  with a single parameter ( $\alpha$ ) for each tempo separately determining a trade-off between the neuromuscular ( $nm$ ) and phrasal ( $phr$ ) templates. Subsequently, we performed a least-square fit of  $\alpha(t) = at^t$  and found the resulting model, indicated by the solid red line, and  $1 - \alpha(t)$  is indicated by the blue line.

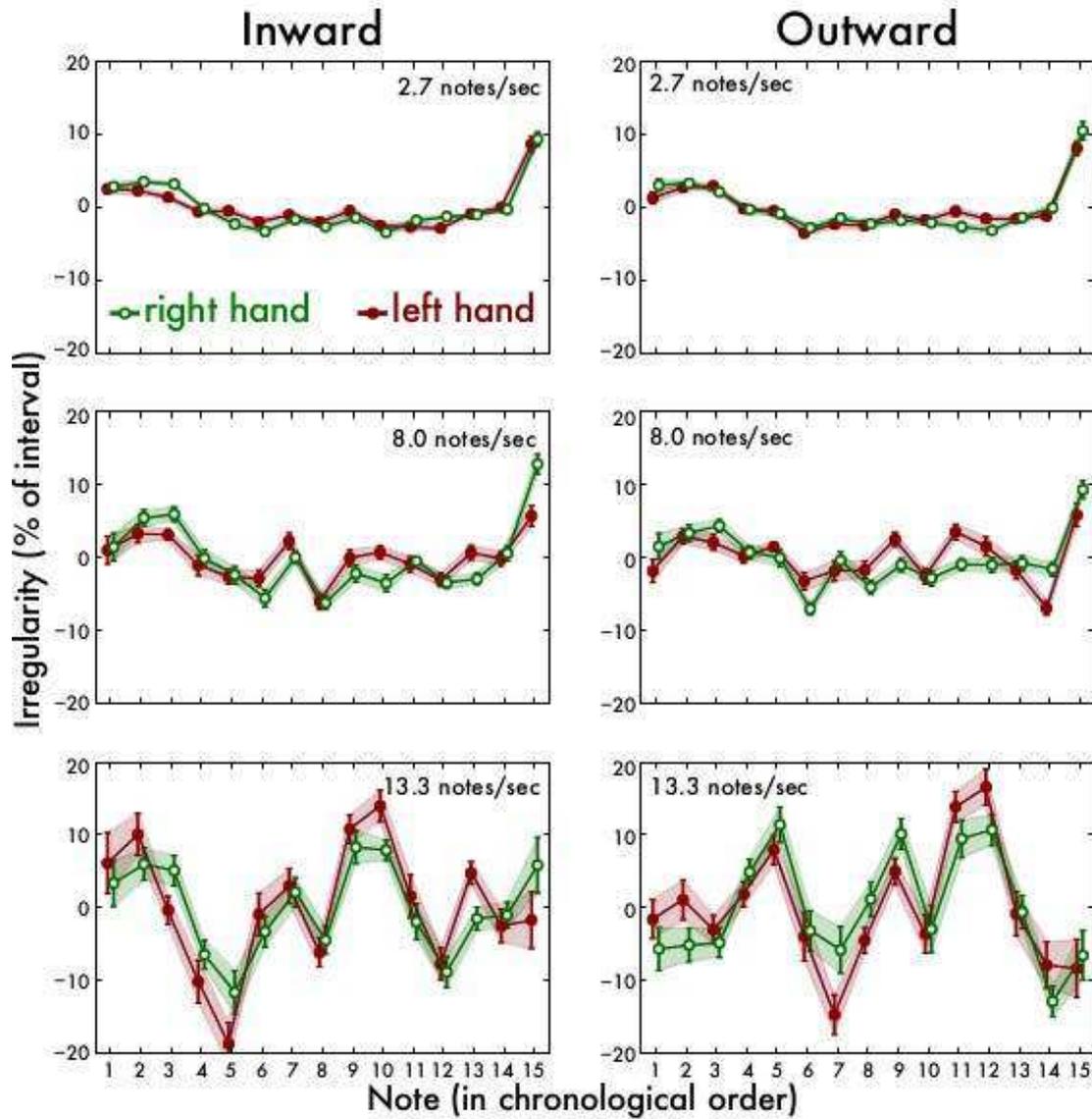


Figure 5. The temporal deviation traces for the inward scale movement (left graphs) and the outward scales movements (right graphs), from the slowest tempo (top), intermediate (middle) to fastest tempo (bottom). For simplicity of presentation, we have omitted the other two measured tempi. The two traces correspond to the left (red) and right (green) hand. Notes are presented in chronological order, i.e. note 1 is the first note played, 2 is the note after that, etc. Note how the traces are similar between the two movements for the slow tempi, but very different for the faster tempi.

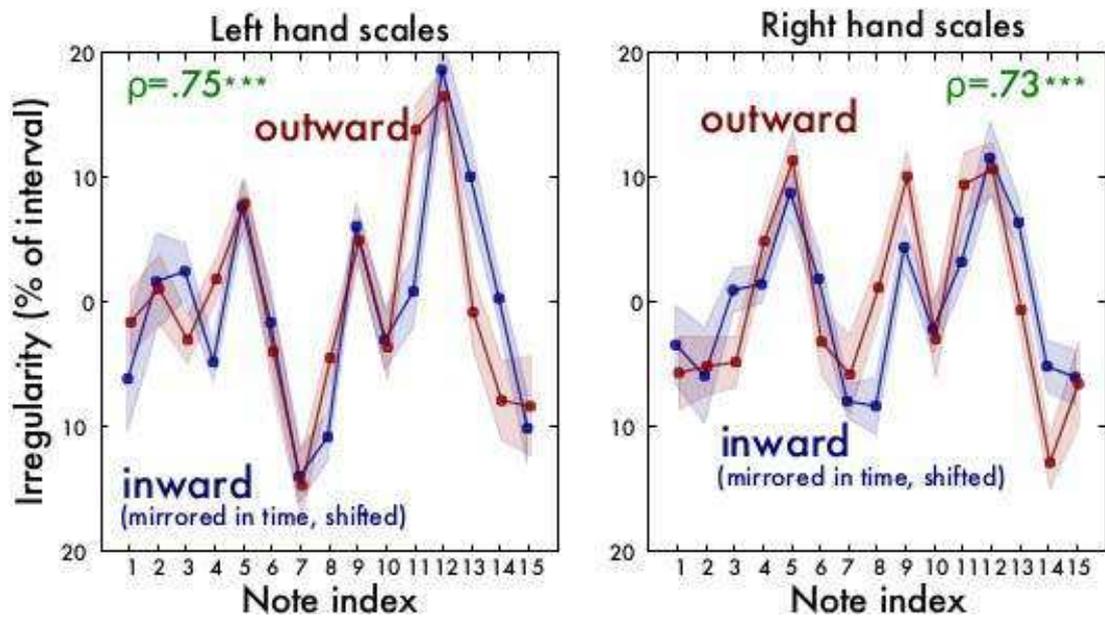


Figure 6. Original outward (red) and re-aligned inward (blue) deviation traces for the fastest tempo. The inward traces are mirrored in time and in lateness, and shifted by one note (see main text for details). Shaded areas indicate the standard error of the mean.

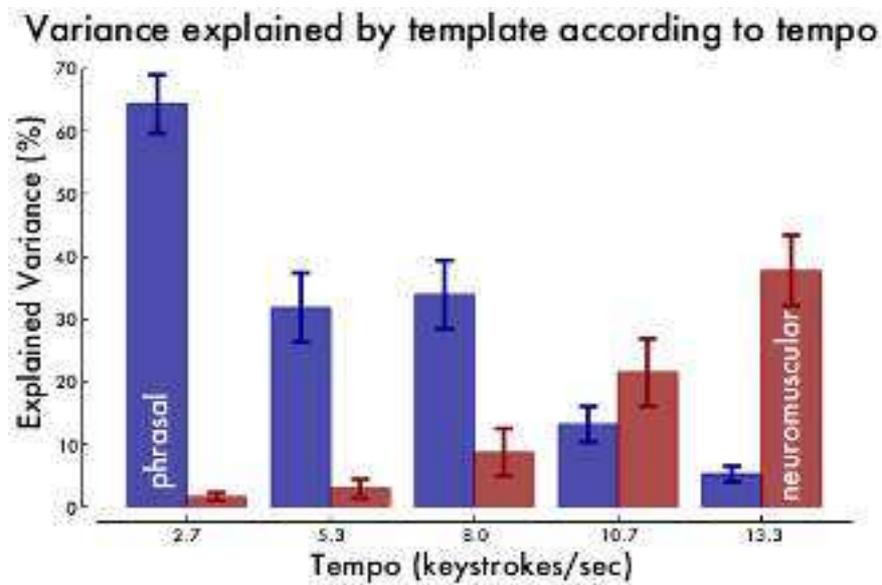


Figure 7. The amount of variance (in percent) that is explained by the phrasal and neuromuscular template according to the playing tempo. A gradual trade-off is evident between the templates at the intermediate tempi with the phrasal template dominating the slow tempi and the neuromuscular template dominating the fast tempi.

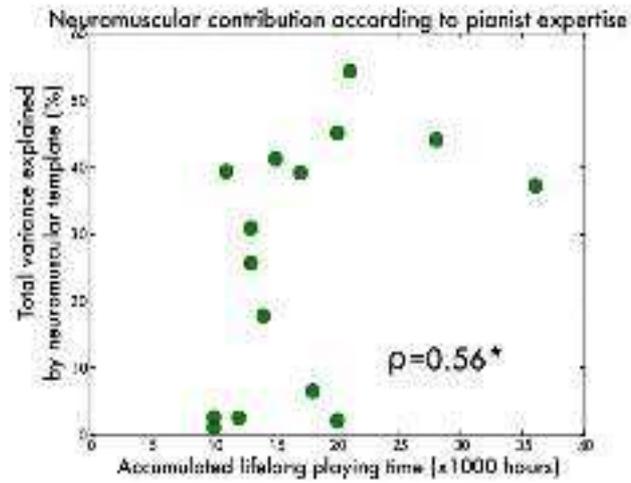


Figure 8. Amount of variance explained (%) by the neuromuscular template for each pianist, as a function of the total lifelong playing time (in hours). Pianists who are more experienced show a greater contribution of the neuromuscular template.

## IV.5. Discussion

---

In this part of the thesis, we asked how expert pianists control timing in musical playing, taking scale playing as an example. In particular, we asked what role auditory feedback plays in motor performance and in learning. In order to answer these questions, the studies presented here have introduced a novel method of analysis of timing deviations away from regularity.

### 5.1 Outline of the irregularity-instability model

The irregularity-instability model is summarised in illustration 12. *Irregularity* refers to the timing deviations that are present on average across trials. For example, some notes are on average early or late by a number of milliseconds. Largely independent of this irregularity is *instability*, which represents the residual variability after the average irregularity trace is taken into account. Put differently, instability measures the trial-to-trial non-correlated variability. The reason we propose to make the distinction between irregularity and instability is the following. Timing deviations that are the result of systematic (non-noise) processes (such as those suggested by the perceptual hypothesis, or due to the neuromuscular apparatus) should not vary between trials. However, non-systematic (noise) processes could vary between different times of day.

Studies IV.2 and IV.3 provided a first presentation of the model and then focused on instability and irregularity, respectively. The dependence of irregularity on tempo is investigated in study IV.4 and variations in instability are the topic of study VIII.1.

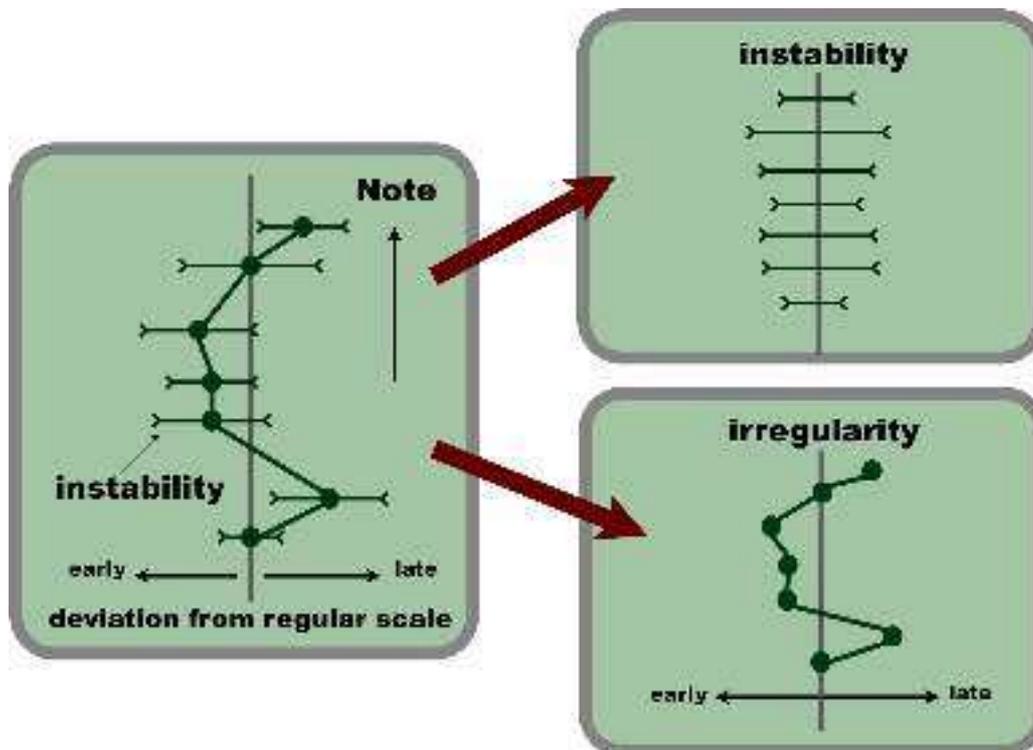


Illustration 12: Summary of the irregularity-instability model of timing in scale playing.

Study IV.3 showed a considerable amount of variation in scale playing that was indeed consistent between trials (irregularity). Put differently, pianists do not time their scales on average regularly, as might have been expected. Instead, pianist-specific timing patterns were found that are robust to the extent that they allowed us to recognise all pianists in adjacent measurement sessions. A portion of pianists was even recognised over a time span of several years. Irregularity, then, is stable over time. Furthermore, there are considerable individual differences in irregularity traces.

In order to understand more deeply what the timing deviations are, study IV.4 went on to model the deviations across a variety of tempi, revealing that the average timing patterns are significantly different across tempi (which conflicts with a generalised motor program hypothesis). At faster tempi, timing is dominated by the difficult movement passages: in our case, the crossover movements for index finger and thumb. At slower

tempi, the timing pattern follows an arc with a slowing down at the beginning and end. This pattern is reminiscent of expressive timing, language phrasing and movements in general.

In light of the robustness of the irregularity traces over time, it is interesting to see that instability is subject to considerable fluctuations. For example, study VIII.1 (included in the appendix for improved readability of the thesis) found that instability was greater in the morning than in the evening for late chronotype pianists. No such difference was found in the irregularity traces. This supports the claim that irregularity traces are stable over time whereas instability varies across the day. Further, study VIII.1 established an important methodological point by showing that the unevenness measure was not able to pick up these fine differences in timing deviation. A possible reason for this is that the measure loses discriminatory power by merging instability with irregularity.

## 5.2 Auditory feedback

What role does auditory feedback play in the control of this movement timing? Experiment III in study IV.3 showed that when pianists played in the absence of auditory feedback, our machine learning algorithm was able to recognise them with virtually the same accuracy as when they played with sound. This shows that the presence or absence of auditory feedback only marginally affected their timing. The finding suggests that the scale playing movement has crystallised into robust motor programs that can be executed in a feed-forward fashion without the need to maintain motor control through feedback mechanisms (Lashley, 1951). This result is further strengthened by results from Exp. I in study IV.3 revealing that timing patterns resulting from the same movement are more similar than timing patterns resulting in the same sounds. If timing performance is based on auditory feedback then the opposite would have been found. Similarly, study IV.4 showed that especially for the faster tempi, timing patterns are governed by movement difficulty: those passages that are more difficult take up more time. The motor system did not compensate for the difficulty in these passages, possibly because the resulting timing deviations were below the perceptual threshold. Notice that we have not measured perceptual sensitivity at this (fast) tempo. A future study could replicate Exp. III of study

IV.3 with fast tempo stimuli in order to investigate whether the timing deviations at that tempo are audible. If so, this would provide evidence against a feedback-based timing mechanism because the brain should have eliminated those deviations.

However, two findings *do* reveal a role for auditory feedback. Firstly, Exp. II in study IV.3 revealed that the timing irregularities as produced by pianists' motor system fall just below the auditory thresholds (although the performing pianists were not the same as the participants in the listening test). This finding is predicted by learning based on a feedback mechanism. Once the motor timing deviations are below the noise level of the feedback stream, the remaining error signal will be noise as well and subsequently not be able to improve the movement anymore. A second finding was that the instability trace (including its octave boundary instability effect) appears to be mirrored by the auditory detection threshold (Exp. III in study IV.2), at least for the set of notes that could be tested in the auditory experiment. What exactly this means is difficult to say, and multiple interpretations are possible (see the paper discussion for details). At the very least, it appears to imply that the auditory and motor system have “chunked” the two-octave scale in a similar way, namely by octave units. The finding provides support for the existence of shared resources (such as chunk formation) between auditory and motor systems.

How can we unite these findings that appear to argue both for and against auditory involvement? We propose that in learning scale playing, learning timing is governed by auditory feedback (Chen, Woollacott, & Moore, 2013; Kajihara, Verdonschot, & Stewart, 2013). Later in learning, as the timing precision falls below the auditory threshold for timing regularity (that is, the scales the pianist produces sound perfectly regular to trained listeners) this process stops. At this point feedback is no longer required, since to the human ear the scales are indistinguishable from perfectly regular scales (study IV.3). Even if feedback had been used, it would not have been able to improve the movements since the irregularities to be corrected are beyond the timing resolution of the auditory system. Instead, the motor system optimises its control by crystallising the motor programs so that they can be executed in a feed-forward fashion.

This account furthermore explains how the octave scale boundary instability effect (study IV.2) is mirrored in auditory perception as well as the fact that scale timing

irregularities straddle participant's auditory perceptual thresholds as predicted by a feedback-based learning model (Exp II in study IV.3). It further accounts for the apparent contradiction that scale playing timing later on appears to virtually cease to rely little on auditory feedback (Exp III in study IV.3).

### 5.3 Future studies

It has been our hope that the studies presented in this part of the thesis may contribute to the investigation of scale playing timing. The investigation in this part of the thesis partly arose from the desire to be able to investigate motor pathologies (such as musician's dystonia) in a more targeted way (study VIII.2). Although this question is not addressed further in this thesis, future research could apply the irregularity-instability analysis to understand the nature of timing deficits in musician's dystonia.

A second promising avenue would be to investigate experimental manipulations in auditory feedback. As described in detail in section 6.4.b of the introduction), a variety of studies have investigated processing of auditory feedback in musical (playing of musical compositions) and non-musical (finger tapping) tasks. However, the influence of auditory feedback has not been directly investigated in piano scale playing. The analytic tools developed here may help to shed light on how musicians control the timing of their playing. Does delayed auditory feedback affect timing irregularity? The interpretation presented here would predict it does not, at least for moderate tempi. The reason is that timing differences appear largely neuromuscular in nature (study IV.3). More advanced manipulations could also be used, such as the method to amplify timing deviations (Exp II in study IV.3). In that study, the method was only used to magnify timing deviations offline during playback of the scale. However, a similar method could be used to amplify timing deviations online as the pianist is playing (for example by introducing a base delay). It is possible that once timing deviations become audible again, the brain reactivates feedback learning. As a result, such amplification of timing deviations might allow pianists to reduce their irregularity trace, perhaps even nearly to zero.

In sum, we provided a novel method of analysis to understand sources of timing deviation away from regularity in the scale playing of pianistic experts. Irregularity, the systematic deviations that do not vary between trials, (1) shows a surprising inter-individual variability that is nevertheless robust, (2) is not dependent on the presence of auditory feedback, and (3) appears mostly neuromuscular in nature. Instability represents residual variability superimposed on an individual's irregularity trace. Instability (1) captures the formation of motor chunks, (2) is reflected by the auditory sensitivity profile, and (3) can vary across moderate time scales (such as across a single day). That is, although experts show traces indicating that they have learned their skill through auditory feedback, expert production appears largely independent of auditory feedback.



## **V. General Discussion**

This thesis investigated sensorimotor integration by asking whether sensory information can contribute to motor processes. In particular, we investigated how human beings learn to make their movements more regular and what role auditory feedback can play during this learning process. The thesis is based on the hypothesis that the role sensory processes play may depend on the population in question. For this reason, we have investigated three populations: healthy non-musicians, non-musician stroke patients and finally musical experts. The claim defended in this thesis is that auditory feedback can be used in all these populations; furthermore, in most natural situations it actually *is* used; but then *how* it is used varies strikingly between the populations.

The discussion chapters (chapters II.4, III.4 and IV.5) in the individual parts of this thesis have already pointed out the major implications of the studies presented therein. The aim of this part of the thesis is to compare the results of each of these studied populations.

## **V.1. Auditory feedback is implicated in motor learning, but *how* depends on the population**

---

The experimental work reported on in this thesis investigated the manipulation of auditory feedback in various populations. The crucial manipulation in parts II and III was to contrast motor learning under the influence of immediate or jittered auditory feedback. The idea was that if the simple presence of sound would be beneficial for motor learning, then a slightly delayed or jittered presence would not deteriorate motor performance. If, however, the timing of the sound relative to the movement is somehow interpreted and used in the motor learning process (as feedback-based models would predict, see section I.2.1.b), we should find that the learning process is affected.

We found that keystroke-triggered sounds allowed a group of healthy non-musicians to improve in motor sequence tapping regularity (study III.2). A *mute* group who did not receive sounds was not able to improve tapping regularity, similar to a *jitter* group who received sounds but with temporal distortions. The learning that was facilitated by auditory feedback furthermore transferred to a novel sequence.

A stroke patient population was able to improve their movement capacity through learning to play music (study III.2). Surprisingly, patients' learning was not impaired by temporal distortions in feedback but enhanced. Their response was to slow down their musical playing, but they subsequently performed better on transfer tasks (such as the pegboard task or finger tapping regularity).

In musical experts we found that scale playing timing instability traces in study IV.2 mirror auditory sensitivity across the notes in the scale. Furthermore, irregularities in pianists' scale playing are just below the auditory threshold of a representative sample of musicians (study IV.3). Finally, we found that absence of auditory feedback did not influence performance in scale playing (study IV.3).

## **1.1 Disruptions in auditory-motor coupling: when it helps and when it hurts**

How to reconcile the seemingly contradictory results found between the various populations?

### **1.1.a Healthy non-musicians: sound feedback-based learning**

#### **i. Auditory error-based learning**

Musical novices were shown to benefit from auditory feedback and were sensitive to temporal disruptions in this feedback. These findings are in line with findings of sensory feedback being crucial for motor learning (Fine & Thoroughman, 2007; Robinson, Noto, & Bevans, 2003). For example, songbirds use auditory feedback, as their learning is affected when pitch noise is added (Sober & Brainard, 2009, 2012). The behaviour of the participants in our study can be explained in roughly the following way. On each trial, they perceived the irregularities in the train of sounds, considering them as a reflection of their movements, i.e. as a feedback signal (see section I.2.1.b). On the next trial, they adjusted for this error by adapting their motor programs. This feedback-based learning model can readily explain why they were not able to learn when their sensory signal was noisy (Sober

& Brainard, 2012; Wei & Körding, 2009) as was the case in our jitter group (study III.2). The noise in the signal would on one trial make them adjust their movements in one direction, and on another trial in another direction. In this way, they would not be able to converge to a regular performance as the normal feedback group did.

The importance of this finding is that it nuances our view of how sound influences movement. We show that it is not the mere presence of sound that improves motor performance, because the jitter group also received sound. Instead, the timing between the movements and the sounds is critical, as is predicted by a feedback-based learning model (Shadmehr, Smith, & Krakauer, 2010).

## ii. Critical time window of integration

An alternative explanation is that a critical time window exists around movements. If a sensory event falls within this time window, it can influence the movement. The idea would be that once sound-to-movement associations are formed (after training; see section I.5.4.e), then the occurrence of that sound would evoke motor activity (see section I.5.4)<sup>5</sup>. This motor activity, in turn, would benefit movement. This idea seems to be supported by findings of audio-spinal facilitation (see section I.6.2.a). Several studies found that for a time window between 50 and 250 msec after a sound, movement is facilitated (C. M. Davis & Beaton, 1968; M. Davis, Gendelman, Tischler, & Gendelman, 1982; Paltsev & Elner, 1967). At this point, we cannot rule out this explanation. However, the account leaves several issues open. First, the sensory effect comes too late. The reason is that once the sensory effect occurs, the movement has already taken place, which means that it can no longer benefit from the auditory-induced motor activity (or that the mere anticipation of sound leads to motor co-activation). Therefore, the claim must be that the sound generates motor facilitation that is spread out over longer time spans (i.e., the sensory effect of one keystroke facilitating the next keystroke). However, this contradicts our finding that the jittered group does not benefit from the sound. This reasoning probably explains why researchers thus far have not made a link between sensory feedback and motor co-activation (Hickok, Houde, & Rong, 2011). Secondly, the account does not make clear

---

5 Possibly, the mere anticipation of the auditory effect would evoke motor activity. However, such motor activity has thus far not been found experimentally. The studies reported in section I.5.4 reported motor activity resulting from actually occurring sound only, not sound that was merely anticipated.

how participants would be able to learn if they are not using sensory feedback. Motor learning is thought to proceed through feedback (whether it be proprioceptive, auditory or visual) (Shadmehr & Krakauer, 2008; Shadmehr et al., 2010). If no feedback is present, a system cannot know what it is doing wrong or right and can therefore not learn to improve.

In sum, the finding that sounds benefit movement when they occur simultaneously (rather than with a jittered delay) are in line with the predictions of an auditory feedback learning model. Potentially, explanations such as a critical time window around the movement could provide alternative accounts.

### **1.1.b Stroke patients: learning not to rely on sound feedback**

In our rehabilitation study with motor stroke patients (part III) we found a different set of results from those obtained in healthy non-musicians. Stroke patients with normal or jittered feedback performed comparably in mood and motor tests, except for a pegboard test and finger tapping regularity in which the jitter group showed a benefit. The findings contradict our expectation, based on a healthy population (part II), that jittered feedback would impair performance.

Further, we found that what was learned through the help of auditory feedback would transfer to other tasks. In the case of patients, the improvements in motor control obtained during music learning (study III.2) transferred to performance improvement in the Nine-Hole Pegboard Test as well as finger tapping. Notice that in those transfer tasks, there was no auditory feedback. No tones were supplied for finger taps or for contact with pegs, but of course some subtle sounds could still have been generated while patients did the tests.

The transfer tasks (especially the Nine Hole Pegboard Test) are very different from the music playing that patients trained. Clearly, achieving this transfer was from the start the aim of the music training. We did not travel to the rehabilitation clinic to turn the patients into expert musicians. Rather, music intervention was used as a tool to improve fine motor control in daily activities, which was the domain in which patients suffered deficits we hoped to alleviate. On the theoretical level, such transfer effects are furthermore interesting because performance differences in such transfer tasks – in which both groups

of participants are in the same condition – are solely attributable to transfer from the feedback-induced learning setting.

**i. Conduction aphasia and stuttering: a parallel to our findings**

The situation is reminiscent of two patient populations suffering from other neurological diseases than motor stroke (These populations were initially introduced in section I.6.3.b). We will base our proposed explanation on these populations. The first group are conduction aphasia patients. These patients struggle to literally repeat speech they hear (Buchsbaum et al., 2011; Fridriksson et al., 2010). Interestingly, they exhibit a decreased sensitivity to normally disruptive effects of delayed auditory feedback (Boller, Vrtunski, Kim, & Mack, 1978). Sometimes, patients even show an improved performance in a delayed condition (Boller & Marcie, 1978). The second group of patients are stutterers. When some of these patients hear their own voices played back with delays their fluency is improved (Martin & Haroldson, 1979). In the case of stutterers the explanation for this phenomenon is the following. Computational models of speech integration can reproduce the effects of stuttering if feedback integration is noisy or when the system relies too much on feedback (Civier, Tasko, & Guenther, 2010; Max, Guenther, Gracco, Ghosh, & Wallace, 2004). Delayed feedback prevents this faulty feedback integration from affecting on-line speech production (Hickok et al., 2011).

**ii. Noisy feedback integration**

Here, we advance the hypothesis that stroke patients may suffer from a similar problem. Our explanation has the following two parts.

Firstly, we hypothesise that stroke patients exhibit faulty auditory-motor integration functioning. This is in line with the proposal advanced by Rodriguez-Fornells et al. (2012) based on the observation that patients show abnormal auditory-motor co-activation prior to therapy relative to non-musician controls (see also Rojo et al., 2011). The novelty of our explanation, however, is that we propose that this faulty auditory-motor integration circuit leads to disruptions in motor performance, as was proposed to be the case in stutterers or conduction aphasia (as discussed in the previous section). In other words, it becomes counterproductive for patients to rely on auditory feedback (Civier et al., 2010; Hickok et

al., 2011; Max et al., 2004). Notice that this is not an explanation of the entire range of motor impairments that stroke patients exhibit. Many of the movements in which they are impaired do not generate sound, and therefore their impairment cannot be explained by faulty auditory-motor integration. Rather, we propose that faulty auditory-motor integration is one of the deficits in these stroke patients which additionally include purely motor deficits.

Secondly, we propose that the noisy feedback causes patients to not integrate auditory feedback and instead rely on other sources of feedback such as visual or proprioception. In a way, the noise in the feedback signals the patient's brain not to use it or to rely on it less. This is in fact predicted by sensorimotor and multisensory integration models. These show that the brain is sensitive to how noisy feedback signals are and weighs these signals inversely by the corresponding noise estimate (Ernst & Banks, 2002; Hillis, Ernst, Banks, & Landy, 2002; Kording & Wolpert, 2004; van Beers, Sittig, & Gon, 1999). Indeed, songbirds are shown to discard auditory pitch feedback error signals if the signal is too noisy (Sober & Brainard, 2012). In other words, in the stroke study delay group (study III.2), the patient's brain realises that the sound feedback is noisy and therefore does not use it as a feedback signal. Instead, it can use visual, proprioceptive or tactile information (Mulder & Hulstijn, 1985). It is possible for the brain to realise that auditory feedback is noisy even though the amount of noise is below the perceptual threshold (as established by the auditory-motor delay detection task). This is not unlikely, given that the noise could affect learning in the healthy individuals (study II.3) although it was not noticeable. As we discussed in section III.1.2, evidence that proprioception is related to clinical outcome is not conclusive. Many studies have found that for those who have proprioceptive deficits, motor outcome is the same as those who have intact proprioception (Rand, (Tamar) Weiss, & Gottlieb, 1999). And even among patients with impaired proprioception, visual and tactile modalities may still provide feedback for learning (Bernier, Chua, Bard, & Franks, 2006; Sarlegna, Malfait, Bringoux, Bourdin, & Vercher, 2010).

How does this explanation account for the effect of jittered feedback in healthy individuals (part II)? The theory outlined in this section would predict that the healthy participants in the jitter group show impaired learning relative the normal group. In the jitter case, the brain also realises the auditory feedback is noisy, but is not able to find an

alternative feedback signal, since we have experimentally eliminated visual feedback and reduced tactile information. The only non-auditory feedback participants can use here is proprioceptive. This explains why some learning was found (such as in speed and accuracy) but the learning was less relative to a group who could integrate sound.

The explanation presented here is supported by the finding that patients with jittered feedback slowed down their music production movements (study III.2). This parallels the stutterers who slow down production (Civier et al., 2010) as well as musicians' slowing down in response to delayed auditory feedback (Pfordresher & Palmer, 2002).

Note that our explanation relies on the postulate that stroke patients exhibit faulty auditory-motor integration circuitry (as argued by Rodriguez-Fornells et al, 2012). If this were not so, we could not account for why the normal group in study III.2 showed less rehabilitation benefit than the delayed feedback group. The normal group had all the same information and in addition, their auditory feedback signal contained more information about their movement (see section I.6). Therefore they should have done at least as well as the delayed feedback group, unless their auditory-motor integration is faulty.

How does our theory account for the fact that patients *did* improve in previous studies investigating music-supported stroke therapy (Amengual et al., 2013; Rojo et al., 2011; Schneider, Schönle, Altenmüller, & Münte, 2007)? If we assume that at least a portion of these patients suffers from auditory-motor integration deficits, how were they able to improve their movement through music therapy? Our explanation for this fact as well as for the improvement in normal groups in our studies III.2 and III.3 is the following. They were able to use visual, proprioceptive and tactile feedback as patients do in any stroke therapies, allowing a majority of them to improve their motor capacities (Dimyan & Cohen, 2011; Liepert, Bauder, Miltner, Taub, & Weiller, 2000; Taub, 2012). In constraint-induced training, the gold standard of medical care, patients improve without auditory feedback. Therefore, in a therapeutic setting such as in our studies, where patients train in making repetitive finger movements, they are able to improve. In addition, patients undergoing music-supported therapy benefit from a morale boost due to the fact that they learn to play music and songs that they know (see section III.1.3.c). The psychological impact of this morale boost, together with the power of music to motivate and move us, might have led to a more active patient participation in music-supported therapy than in other therapies.

A criticism to our explanation could be the following. We argue that auditory-motor integration may be flawed in stroke patients. But why would proprioceptive-motor integration, or visuo-motor integration, or tactile-motor integration be functional? First of all, notice that this criticism can be equally applied to the explanation advanced in (Altenmüller, Marco-Pallares, Münte, & Schneider, 2009), namely that failing proprioceptive feedback was replaced by functioning auditory feedback. The same question applies here: why would proprioceptive feedback be impaired but not auditory feedback? Likely the neural infrastructure for the motor integration of these various modalities is different. For example, auditory-motor integration appears to be intricately linked to the area Spt around the temporal-parietal juncture (Fridriksson et al., 2010; Hickok et al., 2011). Proprioceptive integration is usually thought to happen elsewhere, such as, perhaps, in subcortical structures (Wagle Shukla et al., 2013). However, we feel that this criticism is important and should not be discarded too quickly. Future research may further clarify the question of whether feedback integration of various modalities has different neural infrastructure.

In this account, what is the role of auditory-motor co-activation? Auditory-motor co-activation has been interpreted as the neural substrate responsible for the motor improvement in these stroke patients (Altenmüller et al., 2009; Amengual et al., 2013). However, we argue that there is no direct evidence that motor co-activation during passive music perception is critically involved in motor rehabilitation (see section III.1.3.c). Rather, our interpretation would speak for auditory-motor integration as a side-effect of musical training due to repeated association between auditory and motor events (see I.7 for a discussion of criticality in sensorimotor integration). Future studies could aim to clarify this issue further by investigating the amount of auditory-motor co-activation in a normal and delayed piano therapy group. One prediction that can be derived from the present data is that the patients in the delay group will show less motor co-activation. Furthermore, correlating the amount of motor co-activation with motor rehabilitation benefit could help clarify whether the two are related.

The explanation we provide here might be understood as taking away the promise of sensorimotor integration to boost motor rehabilitation through music. However, notice that in our account, motor learning is still based on feedback. The difference is that the feedback is not auditory, as was hypothesised, but tactile, proprioceptive or visual. In other

words, we argue that sensorimotor integration is the driving force behind music-supported therapy. Our view is that the present studies provide a method for how to use the power of music (to increase patient participation and motivation and therefore therapy effectiveness) while at the same time avoiding aspects of it (such as auditory-motor feedback integration) which might be counterproductive in rehabilitation. Using this more nuanced view, music rehabilitation can be made more effective. Our manipulation of jittered auditory feedback provides a uniquely skillful method to offer a musical rehabilitation paradigm in such a way that it maximally benefits patients.

### **iii. Critical time window of integration**

In section 1.1.a above we introduced the alternative explanation that auditory events could benefit motor learning if they fall within a critical time window. How could this account explain the findings that patients benefit from jittered delay in the auditory feedback? We would need to hypothesise that patients somehow did obtain a larger window of integration in study III.2. As a result of such larger critical window, the patients could benefit from the sound even when it was delayed. In this explanation (and with no other factors at play) we expect that a constant delayed stimulus would benefit patients as much as the jittered delay did in study III.2. On the other hand, the explanation based on noisy feedback integration predicts that patients perform as the normal group did in study III.2, unless the feedback is delayed so much that it is no longer integrated (Sober & Brainard, 2012).

A possible criticism of this explanation is that it remains unclear why neurologically normal individuals do not show such a large integration window. In healthy systems, too, this account predicts it will be beneficial to have a large critical window. By having a larger critical integration window, each movement could benefit from a larger temporal range of sounds. However, the study on healthy non-musicians (study II.3) shows this population does not benefit from jitter delayed feedback, suggesting that they maintained a small integration window even though that deteriorated their performance. This criticism could be countered by hypothesising that it is a feature of damaged auditory-motor systems that the critical window between movement and sound widens. This leaves the question open as to why it is so small in healthy individuals. Perhaps a smaller window of integration allows

for more movement precision. To our knowledge, no studies to date support this idea.

Another possible criticism is that this account does not explain why the delay group performed better than the normal feedback group in study III.2. By hypothesis, both groups should have a larger temporal integration window (as part of the reasoning presented above). Therefore the sound should fall within this temporal integration window both when it is immediate and when it is delayed. In this case, one cannot explain why jittered feedback appeared to improve rehabilitation. Therefore, additional mechanisms would need to be postulated to explain the potentially beneficial effect of jittered delay in auditory feedback.

#### **iv. Implications**

Previously, it was hypothesised that auditory feedback might supplant proprioceptive feedback in music-supported therapy (Altenmüller et al., 2009). Therefore, one predicts that those showing greater proprioceptive deficit should show greater benefit of music therapy. In light of the current explanation of the data, these predictions will invert. We hypothesise that patients with greater proprioceptive deficits will be more likely to rely on auditory feedback, which in turn may be counterproductive (as argued in the previous sections). Therefore functional outcome would actually be *less* in these patients. Future studies could directly test this by including proprioceptive tests (as introduced in section III.1.2) in the testing battery of music-supported intervention study.

#### **v. Delay vs. jitter**

The experimental evidence presented in this thesis is not able to distinguish whether the fact that the feedback was temporally jittered (noisy), rather than delayed, that affected learning performance. This arises from the fact that it is very difficult to introduce temporal noise into auditory feedback without delaying it, because this would require the sound to sometimes precede movements. But how does a researcher know when the keystroke will happen in time? Some researchers have been able to circumvent this problem by asking participants to respond to a visual cue by pressing a key (Stetson, Cui, Montague, & Eagleman, 2006). By measuring the average response times, the researchers were then able to introduce a sound before they predicted the keystroke to happen. In our case, such a

manipulation was not feasible because our patient population already struggled to produce the keystrokes in the first place, let alone to do so in response to a cue.

Investigating the influence of delayed auditory feedback (or delayed sensory feedback for that matter) on motor learning has, to the best of our knowledge, no precedent. As a result, we cannot rely on empirical information to address the question. Previous studies in speech reported disruptive effects of feedback delays (Stuart, Kalinowski, Rastatter, & Lynch, 2002; Yates, 1963) and similarly in music (Pfordresher, 2003; Pfordresher & Palmer, 2002). But these studies only investigated on-line performance and not learning. Furthermore, they investigated only experts (since all humans are experts at speech) while we investigated learning in novices. In the case of speech, it would be obviously problematic to expose people to long-term delayed feedback if this may affect their speech for indeterminate periods of time in the future.

#### **vi. Limitations**

The comparisons between healthy participants and stroke patients presented here have some limitations. We have not tested both populations on exactly the same test. The musical materials that stroke patients are trained on are too easy for healthy individuals and we expect their performance would therefore be at ceiling from the start.

Another difference is that our healthy non-musician group (in part II) did not train on musical materials like the stroke patients did. In the case of the stroke patients, we wanted to use musical materials because this was thought to be highly motivating (see section III.1.3.c). In the case of the healthy non-musicians, we decided to eliminate the differences in tones between the keys because these would have caused the different training sequences to result in different tonal contours, which might also induce tapping irregularities.

Furthermore, in the case of the stroke patients we focus on their performance in transfer tasks (such as finger tapping or the pegboard test) whereas in the case of healthy non-musicians our investigation focused on their performance within the training paradigm. However, we also investigated transfer by comparing a primary and secondary sequence, as well as a scale (1234321) tapping task.

### **1.1.c Musical experts: learning through feedback, performance through feedforward**

In part IV of this thesis, we investigated timing performance of musical experts. We found that although musical experts exhibit evidence of learning through auditory feedback, their performance of over-learned sequences such as scales is not affected by an absence thereof. This is in line with our findings in musical novices who benefited from immediate auditory feedback, but showed no impaired performance when the sounds were removed (in the normal group in study II.3).

The findings, taken together, provide evidence that the auditory feedback is used for *learning*, and then is no longer needed to maintain the performance. The findings are different from other studies where participants became dependent on feedback and showed a decline in performance in its absence (Salmoni, Schmidt, & Walter, 1984), the so-called *guidance effect*. As a result, the evidence presented here supports the idea that auditory feedback may differ from other kinds of feedback (such as, for example, visual feedback) in that it tends to not induce guidance (Heitger et al., 2012), but allow people to function independently of it.

## **V.2. Sensorimotor integration: toward a unified approach**

---

The present thesis investigated whether auditory feedback can be critically involved in motor learning processes (in particular when learning regularity of motor sequence production). This question has been debated in the literature investigating motor involvement in perception, but not in the literature investigating perceptual involvement in motor processes. In this way, our aim was to contribute to a unification of sensorimotor integration.

### **2.1 From a heterogeneous account...**

In chapter I we presented a conceptual map of the various ways in which the

auditory and motor systems interact. Sensorimotor integration is a broad and heterogeneous term that has almost as many meanings as there are authors in the field. Researchers complain that those who study speech perception do not interact with those studying speech production (Hickok et al., 2011) and vice versa. This is probably even more true between the study fields of action, speech and music (chapter I). The novelty of our contribution in part I of this thesis was to provide a systematic framework within which the various uses of sensorimotor integration can be placed. No such generic framework existed to date.

The present thesis shows that auditory feedback has a different effect in different populations. We show that auditory information can be used by the brain to boost motor learning of regularity in healthy individuals. Auditorimotor integration has furthermore been proposed to be an active ingredient responsible for enhanced motor rehabilitation after stroke (Altenmüller et al., 2009; Rodriguez-Fornells et al., 2012). However, study III.2 in this thesis suggests that distortions in auditory-motor coupling could enhance motor rehabilitation. The finding calls into question, if nothing else, the claim that auditory information is *critically involved* in motor performance (and particularly in motor rehabilitation). Rather, it would suggest that auditory information might distract patients from using other sources of feedback (such as proprioception, if the latter is intact). In other words, sensory information is crucial for motor rehabilitation, but the auditory modality might be counterproductive in this population.

## 2.2 ... to a unified account of sensorimotor integration

In light of the fragmented sensorimotor literature, we point out that the object of study, (human) organisms, are one, and the neural resources used in action, speech and music are at the very least greatly overlapping. As a consequence, sensorimotor integration in action, speech and music show important similarities. We have indeed discussed many such similarities (see I.7) and now briefly return to some issues that may provide material for future research.

**i. Integrating remarks: action**

The mirror neuron literature focuses on mirror neurons in areas that are, generally, agreed to be part of the motor system. The critical finding is that these neurons respond to action observation too. However, the converse appears not to occur: no neurons with (motor) mirror properties have been found in canonical sensory areas. This fact is particularly interesting in the light of our current discussion: mirror neurons are involved in one direction of sensory-to-motor transformations. The fact that co-activation studies reveal bi-directional effects (auditory activation upon silent execution, as well as motor activation upon passive listening) may therefore suggest co-activation gives rise to the question why there are no motor “mirror” neurons found in canonical sensory systems.

**ii. Integrating remarks: speech**

As we have seen, studies in perception of speech sounds have revealed that the motor co-activation is somatotopic. A similar finding was obtained in perception of action words, which was distinct for hand, leg and tongue sounds. However, somatotopy of motor activation during finger movement perception could be inverted by training (Catmur, Walsh, & Heyes, 2007). Therefore, a logical question is whether the action word somatotopy is likewise plastic. The experimental paradigm could be as follows. Participants are trained to make a kicking movement in response to the word “lick”, a tongue movement in response to “pick” and a finger movement in response to “kick”. After sufficient training, will the somatotopy of the motor activation evoked by these stimuli be adjusted? If so, it would provide an important piece of evidence for the motor involvement in semantic processing (see section I.5.3.c).

**iii. Integrating remarks: music**

A question that is hardly touched in the sensorimotor literature pertaining to music is to what extent is the motor involvement in music perception *critical*? Various studies reported in this review have investigated facilitatory effects of music perception on movement. However, none have shown the opposite (and quite crucial) prediction of

sensorimotor integration: namely that auditory perception of music should be impaired when motor functioning is impaired.

One way to address this question would be for future studies to replicate the motor disruption paradigms (D'Ausilio et al., 2009) used in speech in the context of professional musicians. For example, TMS could be applied to professional musicians' motor cortex while they listened to musical sounds. An integrative interpretation of sensorimotor interactions predicts that auditory perception of musical sounds should be impaired when motor cortex activity is disrupted.

Secondly, a line of empirical research could investigate processing of degraded musical stimuli. Recall that in speech perception it was observed that motor involvement is enhanced when processing degraded stimuli (section I.5.3.a). Such results would be interesting in the case of rehabilitation approaches that have been proposed to be based on sensorimotor coupling. If degraded stimuli yield enhanced motor activity, should such clinical trials use degraded sounds instead? Further, no study to date investigates a counterpart to the lexical frequency effect in music. Do musically more surprising or infrequent melodies yield enhanced motor activity?

#### **iv. Criticality of sensorimotor integration**

Insight in the exact differences in the use of the notion of sensorimotor integration according to the research community in question will keep us from naively equating things that are called the same with things that are actually the same. For example, researchers have proposed that feedback integration and sensory suppression rely on different neural circuitry (Hain et al., 2000). Similarly, if auditory-motor activation while listening to music is taken to reflect the fact that music makes us want to move (Zatorre, Chen, & Penhune, 2007), then why does hearing speech not make us want to speak (or at least move)? These and many other facts point to important differences between phenomena that are called sensorimotor integration but might be different in nature.

Recognising these differences will be an important first step allowing us go beyond these theoretical and methodological differences and move towards a unified account of sensorimotor integration as it happens in the brain.

What is more important, an understanding of the differences in sensorimotor integration between populations provides an important step to improving our therapies, as has been shown by study III.2 (in comparison with II.3). Simply put, there is a need to understand the effect of auditory feedback in order to be able to choose the most effective application (normal feedback or jittered delay) according to the population in question.

The emerging view presented in the conceptual map of sensorimotor integration (part I) is that during motor actions, sensory systems activate to internally simulate (predict) the results of our motor actions. The generation of the sensory prediction explains the co-activation of the sensory system in an otherwise purely motor task (section I.6.3.a). This prediction allows the brain to then filter out this sensory information from the other influx (such as in sensory suppression, see section I.5.1.b) or use the temporal prediction in a feedback-learning setting (to improve movement) (section 6.1) or align our motor actions to sensation periodically in sensorimotor synchronisation (section I.6.2). Similarly, during perception the motor system is engaged to simulate observed actions by reasoning back from the sensory consequences that were their basis (section I.5.2). This allows mapping acoustic data to phonemic units, participation in conversation and action-perception associations to arise (whose role is debated).

The crucial question in sensorimotor integration, our our view, is to understand what the role is of sensory systems in motor performance, and similarly, what the role is of motor involvement in perception. Previously held views were that the involvement in both ways was *critical* (as Hickok et al. 2011). However, converging evidence from the many fields elaborated on in chapter I makes such positions difficult to hold: the sensorimotor associations are simply too pliable and can be specifically impaired. Therefore, current lines of thought take sensory involvement in motor processing and motor involvement in sensory processing to be modulatory.

Involvement	Description
<b>Critical</b>	The processing in system A is necessary for system B. If system A is damaged, system B cannot function.
<b>Modulatory</b>	System A can influence the processing in system B. Without system A, B can still function but may be impaired.
<b>None</b>	The functioning of system B is not in any way dependent on system A. Any changes in system A do not affect system B.

*Table 1:* Hierarchy of criticality of involvement between two systems (A and B). In our case, A could be the motor system and B the sensory system, or vice versa.

We propose a hierarchy of criticality of sensorimotor integration (Table 1) (although this hierarchy of criticality is general enough to apply to any two processing systems). That is, we propose that there are three levels of motor involvement in sensory processes and, by analogy, of sensory involvement in motor processes. The studies presented in this thesis support the interpretation that auditory sensory information is modulatory to motor learning. We show that participants are able to improve their motor actions in the absence of auditory feedback. This shows that the sensory system is not critically involved in motor learning in these cases. However, the presence or absence of auditory feedback, as well as its timing relative to the movement, did have an important modulatory effect on the performance in the task itself (in the case of healthy nonmusician motor learning) or on transfer tasks (in the case of motor rehabilitation). Further understanding this modulatory influence, we hope, will be the topic of much exciting and fascinating work in this field far into the future.

## **VI. References and indices**

## References

- Aglioti, S. M., and Pazzaglia, M. (2010). Representing actions through their sound. *Exp. Brain Res.* 206, 141–151.
- Aliu, S. O., Houde, J. F., and Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *J. Cogn. Neurosci.* 21, 791–802.
- Altenmüller, E., and Jabusch, H.-C. (2010). Focal dystonia in musicians: phenomenology, pathophysiology, triggering factors, and treatment. *Med. Probl. Perform. Art.* 25, 3–9.
- Altenmüller, E., Marco-Pallarés, J., Münte, T. F., and Schneider, S. (2009). Neural reorganization underlies improvement in stroke-induced motor dysfunction by music-supported therapy. *Ann. N. Y. Acad. Sci.* 1169, 395–405.
- Amengual, J. L., Rojo, N., Veciana de las Heras, M., Marco-Pallarés, J., Grau-Sánchez, J., Schneider, S., Vaquero, L., Juncadella, M., Montero, J., Mohammadi, B., et al. (2013). Sensorimotor Plasticity after Music-Supported Therapy in Chronic Stroke Patients Revealed by Transcranial Magnetic Stimulation. *PLoS ONE* 8, e61883.
- Aschersleben, G., Gehrke, J., and Prinz, W. (2001). Tapping with peripheral nerve block. a role for tactile feedback in the timing of movements. *Exp. Brain Res.* 136, 331–339.
- Avenanti, A., Candidi, M., and Urgesi, C. (2013). Vicarious motor activation during action perception: beyond correlational evidence. *Front. Hum. Neurosci.* 7, 185.
- Aziz-Zadeh, L., Iacoboni, M., Zaidel, E., Wilson, S., and Mazziotta, J. (2004). Left hemisphere motor facilitation in response to manual action sounds. *Eur. J. Neurosci.* 19, 2609–2612.
- Aziz-Zadeh, L., Koski, L., Zaidel, E., Mazziotta, J., and Iacoboni, M. (2006). Lateralization of the Human Mirror Neuron System. *J. Neurosci.* 26, 2964–2970.
- Bangert, M., and Altenmüller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci.* 4, 26–26.
- Bangert, M., Haeusler, U., and Altenmüller, E. (2001). On practice: how the brain connects piano keys and piano sounds. *Ann. N. Y. Acad. Sci.* 930, 425–428.
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H.-J., and Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage* 30, 917–926.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660.
- Basso, A., Casati, G., and Vignolo, L. A. (1977). Phonemic identification defect in aphasia. *Cortex J. Devoted Study Nerv. Syst. Behav.* 13, 85–95.
- Bäss, P., Jacobsen, T., and Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol.* 70, 137–143.
- Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., and Jancke, L. (2007). A network for audio–motor coordination in skilled pianists and non-musicians. *Brain Res.* 1161, 65–78.
- Van Beers, R. J., Sittig, A. C., and Gon, J. J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *J. Neurophysiol.* 81, 1355–1364.
- Bergstrom, I., Seinfeld, S., Arroyo-Palacios, J., Slater, M., and Sanchez-Vives, M. V. (2013). Using music as a signal for biofeedback. *Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol.*
- Bernier, P.-M., Chua, R., Bard, C., and Franks, I. M. (2006). Updating of an internal model without proprioception: a deafferentation study. *Neuroreport* 17, 1421–1425.

- Bever, T. G., and Poeppel, D. (2010). Analysis by synthesis: a (re-) emerging program of research for language and vision. *Biolinguistics* 4, 174–200.
- Bigand, E., and Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition* 100, 100–130.
- Binkofski, F., and Buccino, G. (2004). Motor functions of the Broca’s region. *Brain Lang.* 89, 362–369.
- Bishop, D. V. M. (1988). “Language development after focal brain damage,” in *Language development in exceptional circumstances*, eds. D. V. M. Bishop and K. Mogford (New York: Churchill Livingstone), 203.
- Blakemore, S. J., Wolpert, D. M., and Frith, C. (2000). Why can’t you tickle yourself? *Neuroreport* 11, R11–16.
- Blohm, G., and Crawford, J. D. (2007). Computations for geometrically accurate visually guided reaching in 3-D space. *J. Vis.* 7, 4.1–22.
- Boller, F., and Marcie, P. (1978). Possible role of abnormal auditory feedback in conduction aphasia. *Neuropsychologia* 16, 521–524.
- Boller, F., Vrtunski, P. B., Kim, Y., and Mack, J. L. (1978). Delayed auditory feedback and aphasia. *Cortex* 14, 212–226.
- Boyer, E. O., Babayan, B. M., Bevilacqua, F., Roby-Brami, A., Hannelton, S., and Viaud-Delmon, I. (2013). From ear to hand: the role of the auditory-motor loop in pointing to an auditory source. *Front. Comput. Neurosci.* 7, 26.
- Bradt, J., Dileo, C., and Grocke, D. (2010a). Music interventions for mechanically ventilated patients. *Cochrane Database Syst. Rev. Online*, CD006902.
- Bradt, J., Magee, W. L., Dileo, C., Wheeler, B. L., and McGilloway, E. (2010b). Music therapy for acquired brain injury. *Cochrane Database Syst. Rev. Online*, CD006787.
- Brass, M., Bekkering, H., and Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychol. (Amst.)* 106, 3–22.
- Bridgeman, B., Hendry, D., and Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Res.* 15, 719–722.
- Brinkman, C. (1984). Supplementary motor area of the monkey’s cerebral cortex: short- and long-term deficits after unilateral ablation and the effects of subsequent callosal section. *J. Neurosci.* 4, 918–929.
- Brittain, J.-S., and Brown, P. (2013). Oscillations and the basal ganglia: Motor control and beyond. *Neuroimage*.
- Brown, R. M., and Palmer, C. (2012). Auditory-motor learning influences auditory memory for music. *Mem. Cognit.* 40, 567–578.
- Brozzoli, C., Cardinali, L., Pavani, F., and Farnè, A. (2010). Action-specific remapping of peripersonal space. *Neuropsychologia* 48, 796–802.
- Brozzoli, C., Pavani, F., Urquizar, C., Cardinali, L., and Farnè, A. (2009). Grasping actions remap peripersonal space. *NeuroReport* 20, 913–917.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D’Esposito, M., and Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - an aggregate analysis of lesion and fMRI data. *Brain Lang.* 119, 119–128.
- Buhrmann, T., Paolo, E. A. D., and Barandiaran, X. (2013). A dynamical systems account of sensorimotor contingencies. *Front. Cogn.* 4, 285.
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198.
- Bütefisch, C. M., Davis, B. C., Sawaki, L., Waldvogel, D., Classen, J., Kopylev, L., and Cohen, L. G. (2002). Modulation of use-dependent plasticity by d-amphetamine. *Ann. Neurol.* 51, 59–68.
- Butler, A. J., and James, K. H. (2013). Active learning of novel sound-producing objects: motor reactivation and enhancement of visuo-motor connectivity. *J. Cogn. Neurosci.* 25, 203–218.
- Butler, A. J., James, T. W., and James, K. H. (2011). Enhanced multisensory integration and motor reactivation

- after active motor learning of audiovisual associations. *J. Cogn. Neurosci.* 23, 3515–3528.
- Caetano, G., Jousmäki, V., and Hari, R. (2007). Actor's and observer's primary motor cortices stabilize similarly after seen or heard motor actions. *Proc. Natl. Acad. Sci.* 104, 9058–9062.
- Cai, S., Ghosh, S. S., Guenther, F. H., and Perkell, J. S. (2011). Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *J. Neurosci. Off. J. Soc. Neurosci.* 31, 16483–16490.
- Calvert, G. A., Campbell, R., and Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol. CB* 10, 649–657.
- Calvo-Merino, B., Ehrenberg, S., Leung, D., and Haggard, P. (2010). Experts see it all: configural effects in action observation. *Psychol. Res.* 74, 400–406.
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers. *Cereb. Cortex* 15, 1243–1249.
- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr. Biol. CB* 16, 1905–1910.
- Candidi, M., Urgesi, C., Ionta, S., and Aglioti, S. M. (2008). Virtual lesion of ventral premotor cortex impairs visual perception of biomechanically possible but not impossible actions. *Soc. Neurosci.* 3, 388–400.
- Capek, C. M., Macsweeney, M., Woll, B., Waters, D., McGuire, P. K., David, A. S., Brammer, M. J., and Campbell, R. (2008). Cortical circuits for silent speechreading in deaf and hearing people. *Neuropsychologia* 46, 1233–1241.
- Carey, L. M. (1995). Somatosensory Loss after Stroke. *Crit. Rev. Phys. Rehabil. Med.* 7. Available at: <http://www.dl.begellhouse.com/journals/757fcb0219d89390,58e9038f7231b7dc,7af41ca30975d27f.html> [Accessed June 8, 2013].
- Carey, L. M., Oke, L. E., and Matyas, T. A. (1996). Impaired limb position sense after stroke: A quantitative test for clinical use. *Arch. Phys. Med. Rehabil.* 77, 1271–1278.
- Casile, A., and Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Curr. Biol. CB* 16, 69–74.
- Catmur, C. (2013). Sensorimotor learning and the ontogeny of the mirror neuron system. *Neurosci. Lett.* 540, 21–27.
- Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., and Heyes, C. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. *Eur. J. Neurosci.* 28, 1208–1215.
- Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Curr. Biol.* 17, 1527–1531.
- Cavallo, A., Heyes, C., Becchio, C., Bird, G., and Catmur, C. (2013). Timecourse of mirror and counter-mirror effects measured with transcranial magnetic stimulation. *Soc. Cogn. Affect. Neurosci.* Available at: <http://scan.oxfordjournals.org/content/early/2013/06/25/scan.nst085> [Accessed July 4, 2013].
- Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to Musical Rhythms Recruits Motor Regions of the Brain. *Cereb. Cortex* 18, 2844–2854.
- Chen, J. L., Rae, C., and Watkins, K. E. (2012a). Learning to play a melody: An fMRI study examining the formation of auditory-motor associations. *NeuroImage* 59, 1200–1208.
- Chen, J., Woollacott, M., and Moore, G. P. (2013a). Stochastic aspects of motor behavior and their dependence on auditory feedback in experienced cellists. *Front. Hum. Neurosci.* 7, 419.
- Chen, Z., Chen, X., Liu, P., Huang, D., and Liu, H. (2012b). Effect of temporal predictability on the neural processing of self-triggered auditory stimulation during vocalization. *BMC Neurosci.* 13, 55.
- Chen, Z., Jones, J. A., Liu, P., Li, W., Huang, D., and Liu, H. (2013b). Dynamics of Vocalization-Induced Modulation of Auditory Cortical Activity at Mid-utterance. *PLoS One* 8, e60039.
- Chollet, F., Tardy, J., Albucher, J.-F., Thalamas, C., Berard, E., Lamy, C., Bejot, Y., Deltour, S., Jaillard, A., Niclot,

- P., et al. (2011). Fluoxetine for motor recovery after acute ischaemic stroke (FLAME): a randomised placebo-controlled trial. *Lancet Neurol.* 10, 123–130.
- Civier, O., Tasko, S. M., and Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *J. Fluen. Disord.* 35, 246–279.
- Cleeremans, A., Destrebecqz, A., and Boyer, M. (1998). Implicit learning: news from the front. *Trends Cogn. Sci.* 2, 406–416.
- Conde, V., Altenmüller, E., Villringer, A., and Ragert, P. (2012). Task-irrelevant auditory feedback facilitates motor performance in musicians. *Front. Audit. Cogn. Neurosci.* 3, 146.
- Conforto, A. B., Kaelin-Lang, A., and Cohen, L. G. (2002). Increase in hand muscle strength of stroke patients after somatosensory stimulation. *Ann. Neurol.* 51, 122–125.
- Connell, L., Lincoln, N., and Radford, K. (2008). Somatosensory impairment after stroke: frequency of different deficits and their recovery. *Clin. Rehabil.* 22, 758–767.
- Corbetta, D., Sirtori, V., Moja, L., and Gatti, R. (2010). Constraint-induced movement therapy in stroke patients: systematic review and meta-analysis. *Eur. J. Phys. Rehabil. Med.* 46, 537–544.
- Costello, J. T., and Donnelly, A. E. (2011). Effects of cold water immersion on knee joint position sense in healthy volunteers. *J. Sports Sci.* 29, 449–456.
- Cramer, S. C. (2008). Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Ann. Neurol.* 63, 272–287.
- Crapse, T. B., and Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.* 9, 587–600.
- Creem, S. H., and Proffitt, D. R. (2001). Grasping objects by their handles: a necessary interaction between cognition and action. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 218–228.
- Cressman, E. K., and Henriques, D. Y. P. (2009). Sensory Recalibration of Hand Position Following Visuomotor Adaptation. *J. Neurophysiol.* 102, 3505–3518.
- Crinion, J. T., Warburton, E. A., Lambon-Ralph, M. A., Howard, D., and Wise, R. J. S. (2006). Listening to narrative speech after aphasic stroke: the role of the left anterior temporal lobe. *Cereb. Cortex New York N 1991* 16, 1116–1125.
- Cross, E. S., Hamilton, A. F. de C., and Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *Neuroimage* 31, 1257–1267.
- D’Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., and Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur. J. Neurosci.* 24, 955–958.
- D’Ausilio, A., Brunetti, R., Delogu, F., Santonico, C., and Belardinelli, M. O. (2009a). How and when auditory action effects impair motor performance. *Exp. Brain Res.* 201, 323–330.
- D’Ausilio, a, Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009b). The Motor Somatotopy of Speech Perception. *Curr. Biol.* 19, 381–385.
- Davis, C. M., and Beaton, R. D. (1968). Facilitation and adaptation of the human quadriceps stretch reflex produced by auditory stimulation. *J. Comp. Physiol. Psychol.* 66, 483–487.
- Davis, J. I., Benforado, A., Esrock, E., Turner, A., Dalton, R. C., van Noorden, L., and Leman, M. (2012). Four applications of embodied cognition. *Top. Cogn. Sci.* 4, 786–793.
- Davis, M., Gendelman, D. S., Tischler, M. D., and Gendelman, P. M. (1982). A primary acoustic startle circuit: lesion and stimulation studies. *J. Neurosci.* 2, 791–805.
- Davis, M. H., and Johnsrude, I. S. (2003). Hierarchical Processing in Spoken Language Comprehension. *J. Neurosci.* 23, 3423–3431.
- Deecke, L., Scheid, D.-P. P., and Kornhuber, H. H. (1969). Distribution of readiness potential, pre-motion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements. *Exp. Brain Res.* 7, 158–168.

- Demairé, C., Honoré, J., and Coquery, J. M. (1984). Effects of ballistic and tracking movements on spinal proprioceptive and cutaneous pathways in man. *Prep. States Process. Erlbaum Assoc. Hillsdale New Jersey*, 201–216.
- Dennett, D. C. (1993). *Consciousness Explained*. Penguin Adult.
- Deroost, N., and Soetens, E. (2006). Perceptual or motor learning in SRT tasks with complex sequence structures. *Psychol. Res.* 70, 88–102.
- Desantis, A., Weiss, C., Schütz-Bosbach, S., and Waszak, F. (2012). Believing and perceiving: authorship belief modulates sensory attenuation. *PLoS One* 7, e37959.
- Dimyan, M. A., and Cohen, L. G. (2011). Neuroplasticity in the context of motor rehabilitation after stroke. *Nat Rev Neurol* 7, 76–85.
- Ditchburn, R. W., and Ginsborg, B. L. (1952). Vision with a Stabilized Retinal Image. *Nature* 170, 36–37.
- Dominey, P. F. (1997). An Anatomically Structured Sensory-Motor Sequence Learning System Displays Some General Linguistic Capacities. *Brain Lang.* 59, 50–75.
- Dover, G., and Powers, M. E. (2003). Reliability of Joint Position Sense and Force-Reproduction Measures During Internal and External Rotation of the Shoulder. *J. Athl. Train.* 38, 304–310.
- Drake, C. (1991). Perceptual and performed accents in musical sequences. *J. Acoust. Soc. Am.* 90, 2351.
- Drake, C. (1993). Perceptual and performed accents in musical sequences. *Bull. Psychon. Soc.* 31, 107–110.
- Drawing, K., and Aschersleben, G. (2003). Reduced timing variability during bimanual coupling: a role for sensory information. *Q. J. Exp. Psychol. A* 56, 329–350.
- Drawing, K., Hennings, M., and Aschersleben, G. (2002). The contribution of tactile reafference to temporal regularity during bimanual finger tapping. *Psychol. Res.* 66, 60–70.
- Drawing, K., Stenneken, P., Cole, J., Prinz, W., and Aschersleben, G. (2004). Timing of bimanual movements and deafferentation: implications for the role of sensory movement effects. *Exp. Brain Res.* 158, 50–57.
- Eckers, C., Kröger, B. J., Sass, K., and Heim, S. (2013). Neural representation of the sensorimotor speech-action-repository. *Front. Hum. Neurosci.* 7, 121.
- Eliades, S. J., and Wang, X. (2003). Sensory-Motor Interaction in the Primate Auditory Cortex During Self-Initiated Vocalizations. *J. Neurophysiol.* 89, 2194–2207.
- Elsner, B., and Hommel, B. (2001). Effect anticipation and action control. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 229–240.
- Engel, A., Bangert, M., Horbank, D., Hijmans, B. S., Wilkens, K., Keller, P. E., and Keysers, C. (2012). Learning piano melodies in visuo-motor or audio-motor training conditions and the neural correlates of their cross-modal transfer. *NeuroImage*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22484310> [Accessed May 7, 2012].
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.
- Estévez, A. F., Overmier, B., and Fuentes, L. J. (2003). Differential outcomes effect in children: Demonstration and mechanisms. *Learn. Motiv.* 34, 148–167.
- Fadiga, L., Craighero, L., Buccino, G., and Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur. J. Neurosci.* 15, 399–402.
- Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* 73, 2608–2611.
- Fargier, R., Paulignan, Y., Boulenger, V., Monaghan, P., Reboul, A., and Nazir, T. A. (2012). Learning to associate novel words with motor actions: language-induced motor activity following short training. *Cortex J. Devoted Study Nerv. Syst. Behav.* 48, 888–899.
- Feldman, J., and Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain Lang.* 89, 385–392.
- Ferrari, P. F., Rozzi, S., and Fogassi, L. (2005). Mirror neurons responding to observation of actions made with

- tools in monkey ventral premotor cortex. *J. Cogn. Neurosci.* 17, 212–226.
- Fiebach, C. J., and Schubotz, R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? *Cortex J. Devoted Study Nerv. Syst. Behav.* 42, 499–502.
- Fine, M. S., and Thoroughman, K. A. (2007). Trial-by-Trial Transformation of Error Into Sensorimotor Adaptation Changes With Environmental Dynamics. *J. Neurophysiol.* 98, 1392–1404.
- Finney, S. A. (1997). Auditory Feedback and Musical Keyboard Performance. *Music Percept. Interdiscip. J.* 15, 153–174.
- Finney, S. A., and Palmer, C. (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect. *Mem. Cognit.* 31, 51–64.
- Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771.
- Flash, T., and Sejnowski, T. J. (2001). Computational approaches to motor control. *Curr. Opin. Neurobiol.* 11, 655–662.
- Floel, A., Nagorsen, U., Werhahn, K. J., Ravindran, S., Birbaumer, N., Knecht, S., and Cohen, L. G. (2004). Influence of somatosensory input on motor function in patients with chronic stroke. *Ann. Neurol.* 56, 206–212.
- Fogassi, L., and Ferrari, P. F. (2007). Mirror Neurons and the Evolution of Embodied Language. *Curr. Dir. Psychol. Sci.* 16, 136–141.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal Lobe: From Action Organization to Intention Understanding. *Science* 308, 662–667.
- Froni, F., and Semin, G. R. (2013). Comprehension of action negation involves inhibitory simulation. *Front. Hum. Neurosci.* 7, 209.
- Freeman, J. S., Cody, F. W., and Schady, W. (1993). The influence of external timing cues upon the rhythm of voluntary movements in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 56, 1078–1084.
- Friberg, A., and Sundberg, J. (1999). Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners. *J. Acoust. Soc. Am.* 105, 1469–1484.
- Fridriksson, J., Kjartansson, O., Morgan, P. S., Hjaltason, H., Magnusdottir, S., Bonilha, L., and Rorden, C. (2010). Impaired Speech Repetition and Left Parietal Lobe Damage. *J. Neurosci.* 30, 11057–11061.
- Friederici, A. D. (2006). Broca's area and the ventral premotor cortex in language: functional differentiation and specificity. *Cortex J. Devoted Study Nerv. Syst. Behav.* 42, 472–475.
- Fugl-Meyer, A. R., Jääskö, L., Leyman, I., Olsson, S., and Steglind, S. (1975). The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scand. J. Rehabil. Med.* 7, 13–31.
- Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *J. Neurosci.* 32, 1791–1802.
- Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat. Neurosci.* 7, 773–778.
- Furuya, S., and Altenmüller, E. (2013). Flexibility of movement organization in piano performance. *Front. Hum. Neurosci.* 7, 173.
- Galantucci, B., Fowler, C. A., and Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychon. Bull. Rev.* 13, 361–377.
- Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., and Pizzamiglio, L. (2008). A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage* 40, 1274–1286.
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609.
- Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* 2, 493–501.
- Gandevia, S. C., and Phegan, C. M. (1999). Perceptual distortions of the human body image produced by local

- anaesthesia, pain and cutaneous stimulation. *J. Physiol.* 514 ( Pt 2), 609–616.
- Gandevia, S. C., Refshauge, K. M., and Collins, D. F. (2002). “Sensorimotor Control of Movement and Posture,” in *Advances in Experimental Medicine and Biology*. (Springer US), 61–68. Available at: <http://www.springerlink.com/gate2.inist.fr/content/n1274116422w5022/abstract/> [Accessed March 12, 2012].
- Gates, A., and Bradshaw, J. L. (1974). Effects of auditory feedback on a musical performance task. *Percept. Psychophys.* 16, 105–109.
- Gates, A., Bradshaw, J. L., and Nettleton, N. C. (1974). Effect of different delayed auditory feedback intervals on a music performance task. *Percept. Psychophys.* 15, 21–25.
- Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the Somatotopic Auditory Mirror System in Humans. *Curr. Biol.* 16, 1824–1829.
- Gazzola, V., and Keysers, C. (2009). The Observation and Execution of Actions Share Motor and Somatosensory Voxels in all Tested Subjects: Single-Subject Analyses of Unsmoothed fMRI Data. *Cereb. Cortex New York NY* 19, 1239–1255.
- Gazzola, V., Rizzolatti, G., Wicker, B., and Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* 35, 1674–1684.
- Gebel, B., Braun, C., Kaza, E., Altenmüller, E., and Lotze, M. (2013). Instrument specific brain activation in sensorimotor and auditory representation in musicians. *Neuroimage* 74, 37–44.
- Gentilucci, M., Toni, I., Chieffi, S., and Pavesi, G. (1994). The role of proprioception in the control of prehension movements: a kinematic study in a peripherally deafferented patient and in normal subjects. *Exp. Brain Res.* 99, 483–500.
- Gerrits, D. I. H. J. M., and Vendrik, A. J. H. (1970). Simultaneous contrast, filling-in process and information processing in man’s visual system. *Exp. Brain Res.* 11, 411–430.
- Ghez, C., Gordon, J., and Ghilardi, M. F. (1995). Impairments of reaching movements in patients without proprioception. II. Effects of visual information on accuracy. *J. Neurophysiol.* 73, 361–372.
- Gibson, J. J. (1962). Observations on active touch. *Psychol. Rev.* 69, 477–491.
- Gibson, J. J. (1977). The concept of affordances. *Perceiving Act. Knowing*, 67–82.
- Gilbert, P. F., and Thach, W. T. (1977). Purkinje cell activity during motor learning. *Brain Res.* 128, 309–328.
- Del Giudice, M., Manera, V., and Keysers, C. (2009). Programmed to learn? The ontogeny of mirror neurons. *Dev. Sci.* 12, 350–363.
- Glenberg, A. M., and Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex J. Devoted Study Nerv. Syst. Behav.* 48, 905–922.
- Gonzalez, J., Suzuki, H., Natsumi, N., Sekine, M., and Yu, W. (2012). Auditory display as a prosthetic hand sensory feedback for reaching and grasping tasks. *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf.* 2012, 1789–1792.
- Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Grahn, J. A., and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* 19, 893–906.
- Grahn, J. A., and Rowe, J. B. (2009). Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. *J. Neurosci.* 29, 7540–7548.
- Granert, O., Peller, M., Jabusch, H.-C., Altenmüller, E., and Siebner, H. R. (2011). Sensorimotor skills and focal dystonia are linked to putaminal grey-matter volume in pianists. *J. Neurol. Neurosurg. Psychiatry*, 1225–1231.
- Graziano, M. S. A. (1999). Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proc. Natl. Acad. Sci.* 96, 10418–10421.
- Greeno, J. G. (1994). Gibson’s affordances. *Psychol. Rev.* 101, 336–342.

- Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychol. Rev.* 77, 73–99.
- Gutrecht, J. A., Zamani, A. A., and Pandya, D. N. (1992). Lacunar thalamic stroke with pure cerebellar and proprioceptive deficits. *J. Neurol. Neurosurg. Psychiatry* 55, 854–856.
- Haggard, P., Clark, S., and Kalogerias, J. (2002). Voluntary action and conscious awareness. *Nat. Neurosci.* 5, 382–385.
- Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., and Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Exp. Brain Res. Exp. Hirnforsch. Experimentation Cérébrale* 130, 133–141.
- Hallett, M., Shahani, B. T., and Young, R. R. (1975). EMG analysis of patients with cerebellar deficits. *J. Neurol. Neurosurg. Psychiatry* 38, 1163–1169.
- Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., and Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *J. Cogn. Neurosci.* 17, 282–293.
- Haueisen, J., and Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *J. Cogn. Neurosci.* 13, 786–792.
- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298, 1569–1579.
- Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., and Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology* 42, 180–190.
- Heitger, M. H., Ronsse, R., Dhollander, T., Dupont, P., Caeyenberghs, K., and Swinnen, S. P. (2012). Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis. *Neuroimage* 61, 633–650.
- Helmuth, L. L., and Ivry, R. B. (1996). When two hands are better than one: reduced timing variability during bimanual movements. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 278–293.
- Herholz, S. C., and Zatorre, R. J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron* 76, 486–502.
- He, S. Q., Dum, R. P., and Strick, P. L. (1993). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere. *J. Neurosci. Off. J. Soc. Neurosci.* 13, 952–980.
- Heyes, C. (2010). Where do mirror neurons come from? *Neurosci. Biobehav. Rev.* 34, 575–583.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13, 135–145.
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243.
- Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron* 69, 407–422.
- Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Hillis, J. M., Ernst, M. O., Banks, M. S., and Landy, M. S. (2002). Combining Sensory Information: Mandatory Fusion Within, but Not Between, Senses. *Science* 298, 1627–1630.
- Hirayama, K., Fukutake, T., and Kawamura, M. (1999). “Thumb localizing test” for detecting a lesion in the posterior column-medial lemniscal system. *J. Neurol. Sci.* 167, 45–49.
- Hoffmann, J., Sebold, A., and Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial

- reaction time tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 470–482.
- Hogeveen, J., and Obhi, S. S. (2012). Social interaction enhances motor resonance for observed human actions. *J. Neurosci. Off. J. Soc. Neurosci.* 32, 5984–5989.
- Holmes, G. (1939). The Cerebellum of Man. *Brain* 62, 1–30.
- Holst, E. von, and Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften* 37, 464–476.
- Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878; discussion 878–937.
- Horváth, J., Maess, B., Baess, P., and Tóth, A. (2012). Action-sound coincidences suppress evoked responses of the human auditory cortex in eeg and meg. *J. Cogn. Neurosci.* 24, 1919–1931.
- Houde, J. F., Nagarajan, S. S., Sekihara, K., and Merzenich, M. M. (2002). Modulation of the Auditory Cortex during Speech: An MEG Study. *J. Cogn. Neurosci.* 14, 1125–1138.
- Hughes, G., Desantis, A., and Waszak, F. (2013). Attenuation of auditory N1 results from identity-specific action-effect prediction. *Eur. J. Neurosci.* 37, 1152–1158.
- Hyman, J. M., Whitman, J., Emberly, E., Woodward, T. S., and Seamans, J. K. (2013). Action and Outcome Activity State Patterns in the Anterior Cingulate Cortex. *Cereb. Cortex* 23, 1257–1268.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical Mechanisms of Human Imitation. *Science* 286, 2526–2528.
- Inmann, A., and Haugland, M. (2004). Functional evaluation of natural sensory feedback incorporated in a hand grasp neuroprosthesis. *Med. Eng. Phys.* 26, 439–447.
- Iordanescu, L., Grabowecky, M., and Suzuki, S. (2013). Action enhances auditory but not visual temporal sensitivity. *Psychon. Bull. Rev.* 20, 108–114.
- Jabusch, H.-C. (2006). “Movement Analysis in Pianists,” in *Music, motor control and the brain*, eds. E. Altenmüller, M. Wiesendanger, and J. Kesselring (Oxford, New York: Oxford University Press).
- Jabusch, H.-C., Alpers, H., Kopiez, R., Vauth, H., and Altenmüller, E. (2009). The influence of practice on the development of motor skills in pianists: a longitudinal study in a selected motor task. *Hum. Mov. Sci.* 28, 74–84.
- Jabusch, H.-C., Vauth, H., and Altenmüller, E. (2004). Quantification of focal dystonia in pianists using scale analysis. *Mov. Disord.* 19, 171–180.
- Jabusch, H.-C., Yong, R., and Altenmüller, E. (2007). Music-related motor skills in pianists - Predictors of skill acquisition in childhood and of maintenance in adulthood. in *Proceedings of the International Symposium on Performance Science (AEC)*. Available at: <https://jyx.jyu.fi/dspace/handle/123456789/20881> [Accessed July 11, 2011].
- Jacobs, O. L. R. (1974). *Introduction to control theory*. Clarendon Press.
- James, K. H., and Atwood, T. P. (2009). The role of sensorimotor learning in the perception of letter-like forms: tracking the causes of neural specialization for letters. *Cogn. Neuropsychol.* 26, 91–110.
- James, W. (1918). *The Principles of psychology*. Holt.
- Janata, P., and Edwards, W. H. (2013). A novel sonification strategy for auditory display of heart rate and oxygen saturation changes in clinical settings. *Hum. Factors* 55, 356–372.
- Janata, P., and Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat. Neurosci.* 6, 682–687.
- Janata, P., Tomic, S. T., and Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol. Gen.* 141, 54–75.
- Jäncke, L. (2012). The dynamic audio-motor system in pianists. *Ann. N. Y. Acad. Sci.* 1252, 246–252.
- Jeannerod, M. (2001). Neural Simulation of Action: A Unifying Mechanism for Motor Cognition. *NeuroImage* 14, S103–S109.
- Johnson, E. O., Babis, G. C., Soultanis, K. C., and Soucacos, P. N. (2008). Functional neuroanatomy of

- proprioception. *J. Surg. Orthop. Adv.* 17, 159–164.
- Jola, C., Abedian-Amiri, A., Kuppuswamy, A., Pollick, F. E., and Grosbras, M.-H. (2012). Motor Simulation without Motor Expertise: Enhanced Corticospinal Excitability in Visually Experienced Dance Spectators. *PLoS ONE* 7, e33343.
- Kajihara, T., Verdonschot, R. G., and Stewart, L. (2013). Action-perception coupling in violinists. *Front. Hum. Neurosci.* 7, 349.
- Kauramäki, J., Jääskeläinen, I. P., Hari, R., Möttönen, R., Rauschecker, J. P., and Sams, M. (2010). Lipreading and covert speech production similarly modulate human auditory-cortex responses to pure tones. *J. Neurosci. Off. J. Soc. Neurosci.* 30, 1314–1321.
- Keetels, M., and Vroomen, J. (2012). Exposure to delayed visual feedback of the hand changes motor-sensory synchrony perception. *Exp. Brain Res.* 219, 431–440.
- Keller, P. E., and Koch, I. (2008). Action planning in sequential skills: relations to music performance. *Q. J. Exp. Psychol.* 2006 61, 275–291.
- Keller, P. E., and Koch, I. (2006a). Exogenous and endogenous response priming with auditory stimuli. *Adv. Cogn. Psychol.* 2, 269–276.
- Keller, P. E., and Koch, I. (2006b). The planning and execution of short auditory sequences. *Psychon. Bull. Rev.* 13, 711–716.
- Kelso, J. S., Tuller, B., Vatikiotis-Bateson, E., and Fowler, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 812.
- Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., and Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. *Brain Lang.* 107, 16–43.
- Kermadi, Y., Liu, A., Tempini E.M., Ro, I. (1997). Effects of reversible inactivation of the supplementary motor area (SMA) on unimanual grasp and bimanual pull and grasp performance in monkeys. *Somatosens. Mot. Res.* 14, 268–280.
- Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Exp. Brain Res. Exp. Hirnforsch. Expérimentation Cérébrale* 153, 628–636.
- Keysers, C., and Perrett, D. I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends Cogn. Sci.* 8, 501–507.
- Kilner, J. M., Hamilton, A., and Blakemore, S. J. (2007). Interference effect of observed human movement on action is due to velocity profile of biological motion. *Soc. Neurosci.* 2, 158–166.
- Kilner, J. M., Paulignan, Y., and Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Curr. Biol. CB* 13, 522–525.
- Koelsch, S. (2011). Toward a Neural Basis of Music Perception - A Review and Updated Model. *Front. Psychol.* 2. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3114071/> [Accessed April 27, 2013].
- Koelsch, S., Fritz, T., Schulze, K., Alsop, D., and Schlaug, G. (2005). Adults and children processing music: an fMRI study. *Neuroimage* 25, 1068–1076.
- Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., and Friederici, A. D. (2002). Bach speaks: a cortical “language-network” serves the processing of music. *Neuroimage* 17, 956–966.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science* 297, 846–848.
- Kording, K. P., and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247.
- Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.
- Krakauer, J. W. (2006). Motor learning: its relevance to stroke recovery and neurorehabilitation. *Curr. Opin. Neurol.* 19, 84–90.

- Krauzlis, R. J. (2005). The control of voluntary eye movements: new perspectives. *Neurosci. Rev. J. Bringing Neurobiol. Neurol. Psychiatry* 11, 124–137.
- Kujirai, T., Caramia, M. D., Rothwell, J. C., Day, B. L., Thompson, P. D., Ferbert, A., Wroe, S., Asselman, P., and Marsden, C. D. (1993). Corticocortical inhibition in human motor cortex. *J. Physiol.* 471, 501–519.
- Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 387–394.
- Kunde, W., Koch, I., and Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *Q. J. Exp. Psychol. A* 57, 87–106.
- Kurt, S., Fisher, S. E., and Ehret, G. (2012). Foxp2 Mutations Impair Auditory-Motor Association Learning. *PLoS ONE* 7, e33130.
- Kwakkel, G., Kollen, B. J., van der Grond, J., and Prevo, A. J. H. (2003). Probability of Regaining Dexterity in the Flaccid Upper Limb. *Stroke* 34, 2181–2186.
- Lackner, J. R. (1988). Some Proprioceptive Influences on the Perceptual Representation of Body Shape and Orientation. *Brain* 111, 281–297.
- Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired Actions. *J. Neurosci.* 27, 308–314.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar: theoretical prerequisites. Volume 1.* Stanford University Press.
- Langhorne, P., Coupar, F., and Pollock, A. (2009). Motor recovery after stroke: a systematic review. *Lancet Neurol.* 8, 741–754.
- Lappe, C., Herholz, S. C., Trainor, L. J., and Pantev, C. (2008). Cortical Plasticity Induced by Short-Term Unimodal and Multimodal Musical Training. *J. Neurosci.* 28, 9632–9639.
- Lappe, C., Trainor, L. J., Herholz, S. C., and Pantev, C. (2011). Cortical Plasticity Induced by Short-Term Multimodal Musical Rhythm Training. *PLoS ONE* 6, e21493.
- Lashley, K. (1951). “The problem of serial order in behavior,” in *Cerebral mechanisms in behavior: The Hixon Symposium*, ed. L. A. Jeffress (New York: Wiley), 112–146.
- Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., and Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *J. Neurosci. Off. J. Soc. Neurosci.* 29, 2477–2485.
- Lerdahl, F., and Jackendoff, R. (1977). Toward a formal theory of tonal music. *J. Music Theory* 21, 111–171.
- Lhermitte, F. (1983). “utilization Behaviour” and Its Relation to Lesions of the Frontal Lobes. *Brain* 106, 237–255.
- Lhermitte, F., Pillon, B., and Serdaru, M. (1986). Human autonomy and the frontal lobes. Part I: Imitation and utilization behavior: A neuropsychological study of 75 patients. *Ann. Neurol.* 19, 326–334.
- Lieberman, A. M., and Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition* 21, 1–36.
- Liegeois-Chauvel, C., Morin, C., Musolino, A., Bancaud, J., and Chauvel, P. (1989). Evidence for a Contribution of the Auditory Cortex to Audiospinal Facilitation in Man. *Brain* 112, 375–391.
- Liepert, J., Bauder, H., Miltner, W. H. R., Taub, E., and Weiller, C. (2000). Treatment-Induced Cortical Reorganization After Stroke in Humans. *Stroke* 31, 1210–1216.
- Lim, I., van Wegen, E., de Goede, C., Deutekom, M., Nieuwboer, A., Willems, A., Jones, D., Rochester, L., and Kwakkel, G. (2005). Effects of external rhythmical cueing on gait in patients with Parkinson’s disease: a systematic review. *Clin. Rehabil.* 19, 695–713.
- Lodha, N., Coombes, S. A., and Cauraugh, J. H. (2012). Bimanual isometric force control: asymmetry and coordination evidence post stroke. *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.* 123, 787–795.
- Loehr, J. D. (2013). Sensory attenuation for jointly produced action effects. *Front. Psychol.* 4, 172.
- Lotto, A. J., Hickok, G. S., and Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends Cogn. Sci.* 13, 110–114.

- Lotze, M., and Halsband, U. (2006). Motor imagery. *J. Physiol. Paris* 99, 386–395.
- Lotze, M., Scheler, G., Tan, H.-R. M., Braun, C., and Birbaumer, N. (2003). The musician’s brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* 20, 1817–1829.
- De Lucia, M., Camen, C., Clarke, S., and Murray, M. M. (2009). The role of actions in auditory object discrimination. *NeuroImage* 48, 475–485.
- Luo, C., Guo, Z., Lai, Y., Liao, W., Liu, Q., Kendrick, K. M., Yao, D., and Li, H. (2012). Musical Training Induces Functional Plasticity in Perceptual and Motor Networks: Insights from Resting-State fMRI. *PLoS ONE* 7, e36568.
- MacKenzie, C. L., and Van Eerd, D. L. (1990). Rhythmic precision in the performance of piano scales: Motor psychophysics and motor programming. *Atten. Perform.* 13, 375–408.
- MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., Williams, S. C., Woll, B., and Brammer, M. J. (2000). Silent speechreading in the absence of scanner noise: an event-related fMRI study. *Neuroreport* 11, 1729–1733.
- Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70.
- Makin, T. R., Holmes, N. P., Brozzoli, C., Rossetti, Y., and Farnè, A. (2009). Coding of Visual Space during Motor Preparation: Approaching Objects Rapidly Modulate Corticospinal Excitability in Hand-Centered Coordinates. *J. Neurosci.* 29, 11841–11851.
- Malcolm, M. P., Massie, C., and Thaut, M. (2009). Rhythmic auditory-motor entrainment improves hemiparetic arm kinematics during reaching movements: a pilot study. *Top. Stroke Rehabil.* 16, 69–79.
- Manning, F., and Schutz, M. (2013). “Moving to the beat” improves timing perception. *Psychon. Bull. Rev.*, 1–7.
- Margulis, E. H., Mlsna, L. M., Uppunda, A. K., Parrish, T. B., and Wong, P. C. M. (2009). Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Hum. Brain Mapp.* 30, 267–275.
- Martikainen, M. H., Kaneko, K., and Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cereb. Cortex New York N* 1991 15, 299–302.
- Martinez-Conde, S., Otero-Millan, J., and Macknik, S. L. (2013). The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nat. Rev. Neurosci.* 14, 83–96.
- Martin, R., and Haroldson, S. K. (1979). Effects of five experimental treatments of stuttering. *J. Speech Hear. Res.* 22, 132–146.
- Matthews, P. B. C., and Rushworth, G. (1958). The discharge from muscle spindles as an indicator of efferent paralysis by procaine. *J. Physiol.* 140, 421–426.
- Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., and Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemp. Issues Commun. Sci. Disord.* 31, 105–122.
- Mayr, U. (1996). Spatial attention and implicit sequence learning: evidence for independent learning of spatial and nonspatial sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 350–364.
- McCamy, M. B., Otero-Millan, J., Macknik, S. L., Yang, Y., Troncoso, X. G., Baer, S. M., Crook, S. M., and Martinez-Conde, S. (2012). Microsaccadic efficacy and contribution to foveal and peripheral vision. *J. Neurosci. Off. J. Soc. Neurosci.* 32, 9194–9204.
- McNeil, C. J., Butler, J. E., Taylor, J. L., and Gandevia, S. C. (2013). Testing the excitability of human motoneurons. *Front. Hum. Neurosci.* 7, 152.
- Mechsner, F., Kerzel, D., Knoblich, G., and Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature* 414, 69–73.
- Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R., and Thron, A. (2004). Playing piano in the mind--an fMRI study on music imagery and performance in pianists. *Brain Res. Cogn. Brain Res.* 19, 219–228.

- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., and Iacoboni, M. (2007). The Essential Role of Premotor Cortex in Speech Perception. *Curr. Biol.* 17, 1692–1696.
- Ménoret, M., Curie, A., des Portes, V., Nazir, T. A., and Paulignan, Y. (2013). Simultaneous action execution and observation optimise grasping actions. *Exp. Brain Res.*
- Merchant, H., Pérez, O., Zarco, W., and Gámez, J. (2013). Interval Tuning in the Primate Medial Premotor Cortex as a General Timing Mechanism. *J. Neurosci.* 33, 9082–9096.
- Mercier, C., Bertrand, A. M., and Bourbonnais, D. (2004). Differences in the magnitude and direction of forces during a submaximal matching task in hemiparetic subjects. *Exp. Brain Res. Exp. Hirnforsch. Experimentation Cérébrale* 157, 32–42.
- Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804.
- Meyer, M., Zysset, S., von Cramon, D. Y., and Alter, K. (2005). Distinct fMRI responses to laughter, speech, and sounds along the human peri-sylvian cortex. *Brain Res. Cogn. Brain Res.* 24, 291–306.
- Mock, J. R., Foundas, A. L., and Golob, E. J. (2011). Selective influence of auditory distractors on motor cortex excitability. *Neuroreport* 22, 830–833.
- Morin, O., and Grèzes, J. (2008). What is “mirror” in the premotor cortex? A review. *Neurophysiol. Clin. Clin. Neurophysiol.* 38, 189–195.
- Möttönen, R., Dutton, R., and Watkins, K. E. (2013). Auditory-motor processing of speech sounds. *Cereb. Cortex New York N* 1991 23, 1190–1197.
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr. Biol. CB* 20, 750–756.
- Mulder, T., and Hulstijn, W. (1985). Sensory feedback in the learning of a novel motor task. *J. Mot. Behav.* 17, 110–128.
- Müller-Preuss, P., and Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain Res.* 215, 61–76.
- Mutschler, I., Schulze-Bonhage, A., Glauche, V., Demandt, E., Speck, O., and Ball, T. (2007). A Rapid Sound-Action Association Effect in Human Insular Cortex. *PLoS ONE* 2, e259.
- Nasir, S. M., and Ostry, D. J. (2009). Auditory plasticity and speech motor learning. *Proc. Natl. Acad. Sci.* Available at: <http://www.pnas.org/content/early/2009/10/30/0907032106> [Accessed December 25, 2012].
- Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., and Orban, G. A. (2005). Observing Others: Multiple Action Representation in the Frontal Lobe. *Science* 310, 332–336.
- Newell, K. M. (1991). Motor Skill Acquisition. *Annu. Rev. Psychol.* 42, 213–237.
- Nieto-Castanon, A., Guenther, F. H., Perkell, J. S., and Curtin, H. D. (2005). A modeling investigation of articulatory variability and acoustic stability during American English /r/ production. *J. Acoust. Soc. Am.* 117, 3196–3212.
- Nissen, M. J., and Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognit. Psychol.* 19, 1–32.
- Niziolek, C. A., and Guenther, F. H. (2013). Vowel Category Boundaries Enhance Cortical and Behavioral Responses to Speech Feedback Alterations. *J. Neurosci.* 33, 12090–12098.
- Nowak, D. A., Grefkes, C., Dafotakis, M., Eickhoff, S., Küst, J., Karbe, H., and Fink, G. R. (2008). Effects of low-frequency repetitive transcranial magnetic stimulation of the contralesional primary motor cortex on movement kinematics and neural activity in subcortical stroke. *Arch. Neurol.* 65, 741–747.
- Numminen, J., and Curio, G. (1999). Differential effects of overt, covert and replayed speech on vowel-evoked responses of the human auditory cortex. *Neurosci. Lett.* 272, 29–32.
- Nyberg, L., Habib, R., McIntosh, A. R., and Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11120–11124.

- Obleser, J., Scotty, S. K., and Eulitz, C. (2005). Now You Hear It, Now You Don't: Transient Traces of Consonants and their Nonspeech Analogues in the Human Brain. *Cereb. Cortex* 16, 1069–1076.
- Del Olmo, M. F., and Cudeiro, J. (2003). A simple procedure using auditory stimuli to improve movement in Parkinson's disease: a pilot study. *Neurol. Clin. Neurophysiol. NCN* 2003, 1–7.
- Del Olmo, M. F., and Cudeiro, J. (2005). Temporal variability of gait in Parkinson disease: effects of a rehabilitation programme based on rhythmic sound cues. *Parkinsonism Relat. Disord.* 11, 25–33.
- Opitz, B., and Kotz, S. A. (2012). Ventral premotor cortex lesions disrupt learning of sequential grammatical structures. *Cortex J. Devoted Study Nerv. Syst. Behav.* 48, 664–673.
- Oppenheim, G. M., and Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Mem. Cognit.* 38, 1147–1160.
- O'Regan, J. K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–973; discussion 973–1031.
- Oscari, F., Secoli, R., Avanzini, F., Rosati, G., and Reinkensmeyer, D. J. (2012). Substituting auditory for visual feedback to adapt to altered dynamic and kinematic environments during reaching. *Exp. Brain Res. Exp. Hirnforsch. Expérimentation Cérébrale* 221, 33–41.
- Ostry, D. J., Darainy, M., Mattar, A. A. G., Wong, J., and Gribble, P. L. (2010). Somatosensory Plasticity and Motor Learning. *J. Neurosci.* 30, 5384–5393.
- Overy, K. (2012). Making music in a group: synchronization and shared experience. *Ann. N. Y. Acad. Sci.* 1252, 65–68.
- Paillard, J., and Brouchon, M. (1968). Active and passive movements in the calibration of position sense. *Neuropsychol. Spatially Oriented Behav.*, 37–55.
- Palmer, C. (1997). Music Performance. *Annu. Rev. Psychol.* 48, 115–138.
- Paltsev, Y. I., and Elner, A. M. (1967). Change in the functional state of the segmental apparatus of the spinal cord under the influence of sound stimuli and its role in voluntary movement. *Biophysics* 12, 1219–1226.
- Papeo, L., Pascual-Leone, A., and Caramazza, A. (2013). Disrupting the brain to validate hypotheses on the neurobiology of language. *Front. Hum. Neurosci.* 7, 148.
- Papeo, L., Vallesi, A., Isaja, A., and Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS One* 4, e4508.
- Pascual-Leone, A., Nguyet, D., Cohen, L. G., Brasil-Neto, J. P., Cammarota, A., and Hallett, M. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J. Neurophysiol.* 74, 1037–1045.
- Pau, S., Jahn, G., Sakreida, K., Domin, M., and Lotze, M. (2013). Encoding and recall of finger sequences in experienced pianists compared with musically naïve controls: A combined behavioral and functional imaging study. *NeuroImage* 64, 379–387.
- Pavan, A., Skujevskis, M., and Baggio, G. (2013). Motion words selectively modulate direction discrimination sensitivity for threshold motion. *Front. Hum. Neurosci.* 7, 134.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180.
- Penel, A., and Drake, C. (1998). Sources of timing variations in music performance: A psychological segmentation model. *Psychol. Res. Psychol. Forsch.* 61, 12–32.
- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., and Brattico, E. (2011). Music and Emotions in the Brain: Familiarity Matters. *PLoS ONE* 6, e27241.
- Petroni, A., Baguear, F., and Della-Maggiore, V. (2010). Motor Resonance May Originate From Sensorimotor Experience. *J. Neurophysiol.* 104, 1867–1871.
- Pfordresher, P. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 949–964.

- Pfordresher, P., and Dalla Bella, S. (2011). Delayed auditory feedback and movement. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 566–579.
- Pfordresher, P., and Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychol. Res.* 66, 71–79.
- Phillips-Silver, J., and Trainor, L. J. (2005). Feeling the Beat: Movement Influences Infant Rhythm Perception. *Science* 308, 1430–1430.
- Phillips-Silver, J., and Trainor, L. J. (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition* 105, 533–546.
- Plautz, E. J., Milliken, G. W., and Nudo, R. J. (2000). Effects of repetitive motor training on movement representations in adult squirrel monkeys: role of use versus learning. *Neurobiol. Learn. Mem.* 74, 27–55.
- Popescu, M., Otsuka, A., and Ioannides, A. A. (2004). Dynamics of brain activity in motor and frontal cortical areas during music listening: a magnetoencephalographic study. *Neuroimage* 21, 1622–1638.
- Poulet, J. F. A., and Hedwig, B. (2003). Corollary Discharge Inhibition of Ascending Auditory Neurons in the Stridulating Cricket. *J. Neurosci.* 23, 4717–4725.
- Prather, J. F., Peters, S., Nowicki, S., and Mooney, R. (2008). Precise auditory–vocal mirroring in neurons for learned vocal communication. *Nature* 451, 305–310.
- Press, C., Catmur, C., Cook, R., Widmann, H., Heyes, C., and Bird, G. (2012). fMRI Evidence of “Mirror” Responses to Geometric Shapes. *PLoS ONE* 7, e51934.
- Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154.
- Proffitt, D. R. (2006). Embodied Perception and the Economy of Action. *Perspect. Psychol. Sci.* 1, 110–122.
- Proverbio, A. M., Riva, F., and Zani, A. (2009). Observation of static pictures of dynamic actions enhances the activity of movement-related brain areas. *PloS One* 4, e5389.
- Pulvermüller, F., and Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* 11, 351–360.
- Pulvermüller, F., Huss, M., Kherif, F., Martin, F. M. del P., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci.* 103, 7865–7870.
- Rand, D., (Tamar) Weiss, P. L., and Gottlieb, D. (1999). Does Proprioceptive Loss Influence Recovery of the Upper Extremity After Stroke? *Neurorehabil. Neural Repair* 13, 15 –21.
- Reichenbach, A., Thielscher, A., Peer, A., Bulthoff, H. H., and Bresciani, J.-P. (2009). Seeing the hand while reaching speeds up on-line responses to a sudden change in target position. *J. Physiol.* 587, 4605–4616.
- Remillard, G. (2003). Pure perceptual-based sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 581–597.
- Repp, B. H. (1999a). Control of Expressive and Metronomic Timing in Pianists. *J. Mot. Behav.* 31, 145–164.
- Repp, B. H. (1999b). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Percept. Psychophys.* 61, 529–548.
- Repp, B. H. (1992a). Diversity and commonality in music performance: an analysis of timing microstructure in Schumann’s “Träumerei.” *J. Acoust. Soc. Am.* 92, 2546–2568.
- Repp, B. H. (1999c). Effects of auditory feedback deprivation on expressive piano performance. *Music Percept.* 16, 409–438.
- Repp, B. H. (1992b). Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Cognition* 44, 241–281.
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* 12, 969–992.
- Repp, B. H., and Knoblich, G. (2007). Toward a psychophysics of agency: detecting gain and loss of control over auditory action effects. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 469–482.

- Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.* 21, 188–194.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Exp. Brain Res. Exp. Hirnforsch. Expérimentation Cérébrale* 71, 491–507.
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3, 131–141.
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670.
- Robinson, F. R., Noto, C. T., and Bevans, S. E. (2003). Effect of Visual Error Size on Saccade Adaptation in Monkey. *J. Neurophysiol.* 90, 1235–1244.
- Rodriguez-Fornells, A., Rojo, N., Amengual, J. L., Ripollés, P., Altenmüller, E., and Münte, T. F. (2012). The involvement of audio-motor coupling in the music-supported therapy applied to stroke patients. *Ann. N. Y. Acad. Sci.* 1252, 282–293.
- Rojo, N., Amengual, J., Juncadella, M., Rubio, F., Camara, E., Marco-Pallares, J., Schneider, S., Veciana, M., Montero, J., Mohammadi, B., et al. (2011). Music-Supported Therapy induces plasticity in the sensorimotor cortex in chronic stroke: A single-case study using multimodal imaging (fMRI-TMS). *Brain Inj.* Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21561296> [Accessed May 23, 2011].
- Roland, P. E., Larsen, B., Lassen, N. A., and Skinhoj, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *J. Neurophysiol.* 43, 118–136.
- Rolfs, M. (2009). Microsaccades: small steps on a long way. *Vision Res.* 49, 2415–2441.
- Ronsse, R., Puttemans, V., Coxon, J. P., Goble, D. J., Wagemans, J., Wenderoth, N., and Swinnen, S. P. (2011). Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. *Cereb. Cortex New York N 1991* 21, 1283–1294.
- Rosenbaum, D. A. (2009). *Human Motor Control*. Academic Press.
- Rosenkranz, K., and Rothwell, J. C. (2012). Modulation of Proprioceptive Integration in the Motor Cortex Shapes Human Motor Learning. *J. Neurosci.* 32, 9000–9006.
- Rosenzweig, M. R., and Leiman, A. L. (1982). *Physiological psychology*. Lexington, Mass: D.C. Heath.
- Rossignol, S. (1975). Startle responses recorded in the leg of man. *Electroencephalogr. Clin. Neurophysiol.* 39, 389–397.
- Rossignol, S., and Melvill-Jones, G. (1976). Audio-spinal influence in man studied by the H-reflex and its possible role on rhythmic movements synchronized to sound. *Electroencephalogr. Clin. Neurophysiol.* 41, 83–92.
- Roy, A. C., Craighero, L., Fabbri-Destro, M., and Fadiga, L. (2008). Phonological and lexical motor facilitation during speech listening: a transcranial magnetic stimulation study. *J. Physiol. Paris* 102, 101–105.
- De Ruiter, J. P., Mitterer, H., and Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, 515–535.
- Sacco, R. L., Bello, J. A., Traub, R., and Brust, J. C. (1987). Selective proprioceptive loss from a thalamic lacunar stroke. *Stroke J. Cereb. Circ.* 18, 1160–1163.
- Saig, A., Gordon, G., Assa, E., Arieli, A., and Ahissar, E. (2012). Motor-sensory confluence in tactile perception. *J. Neurosci. Off. J. Soc. Neurosci.* 32, 14022–14032.
- Sakreida, K., Scorolli, C., Menz, M. M., Heim, S., Borghi, A. M., and Binkofski, F. (2013). Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. *Front. Hum. Neurosci.* 7, 125.
- Salmoni, A. W., Schmidt, R. A., and Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychol. Bull.* 95, 355–386.
- Santello, M., Flanders, M., and Soechting, J. F. (2002). Patterns of hand motion during grasping and the

- influence of sensory guidance. *J. Neurosci.* 22, 1426–1435.
- Sarlegna, F. R., Malfait, N., Bringoux, L., Bourdin, C., and Vercher, J.-L. (2010). Force-field adaptation without proprioception: can vision be used to model limb dynamics? *Neuropsychologia* 48, 60–67.
- Sato, M., Grabski, K., Garnier, M., Schwartz, J.-L., and Nguyen, N. (2013). Converging toward a common speech code: imitative and perceptuo-motor recalibration processes in speech production. *Front. Cogn. Sci.* 4, 422.
- Saunders, J. A., and Knill, D. C. (2003). Humans use continuous visual feedback from the hand to control fast reaching movements. *Exp. Brain Res. Exp. Hirnforsch. Expérimentation Cérébrale* 152, 341–352.
- Saunders, J. A., and Knill, D. C. (2004). Visual Feedback Control of Hand Movements. *J. Neurosci.* 24, 3223–3234.
- Schabrun, S. M., and Hillier, S. (2009). Evidence for the retraining of sensation after stroke: a systematic review. *Clin. Rehabil.* 23, 27–39.
- Schafer, E. W. P., and Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain Responses. *Science* 181, 175–177.
- Schaffert, N., Mattes, K., and Effenberg, A. O. (2011). An investigation of online acoustic information for elite rowers in on-water training conditions. *J. Hum. SPORT Exerc. - Univ. Alicante* 6, 392–405.
- Schmidt, R. A., and Lee, T. (1988). *Motor Control and Learning, 5E*. Human Kinetics 10%.
- Schmitz, G., Mohammadi, B., Hammer, A., Heldmann, M., Samii, A., Münte, T. F., and Effenberg, A. O. (2013). Observation of sonified movements engages a basal ganglia frontocortical network. *BMC Neurosci.* 14, 32.
- Schneider, S., Münte, T., Rodriguez-Fornells, A., Sailer, M., and Altenmüller, E. (2010). Music-Supported Training is More Efficient than Functional Motor Training for Recovery of Fine Motor Skills in Stroke Patients. *Music Percept.* 27, 271–280.
- Schneider, S., Schönle, P. W., Altenmüller, E., and Münte, T. F. (2007). Using musical instruments to improve motor skill recovery following a stroke. *J. Neurol.* 254, 1339–1346.
- Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Curr. Opin. Neurobiol.* 20, 172–176.
- Schubotz, R. I., and Cramon, D. Y. von (2004). Sequences of Abstract Nonbiological Stimuli Share Ventral Premotor Cortex with Action Observation and Imagery. *J. Neurosci.* 24, 5467–5474.
- Schubotz, R. I., von Cramon, D. Y., and Lohmann, G. (2003). Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *Neuroimage* 20, 173–185.
- Schuch, S., Bayliss, A. P., Klein, C., and Tipper, S. P. (2010). Attention modulates motor system activation during action observation: evidence for inhibitory rebound. *Exp. Brain Res.* 205, 235–249.
- Schuil, K. D. I., and Zwaan, R. A. (2013). Sentential context modulates the involvement of the motor cortex in action language processing: an fMRI study. *Front. Hum. Neurosci.* 7, 100.
- Schwartz, J.-L., Basirat, A., Ménard, L., and Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *J. Neurolinguistics* 25, 336–354.
- Scott, S. K., McGettigan, C., and Eisner, F. (2009). A little more conversation, a little less action--candidate roles for the motor cortex in speech perception. *Nat. Rev. Neurosci.* 10, 295–302.
- Seashore, C. E. (1938). *Psychology of Music*. New York: Dover.
- Shadmehr, R., and Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Exp. Brain Res.* 185, 359–381.
- Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annu. Rev. Neurosci.* 33, 89–108.
- Shaffer, L. H. (1984a). Timing in Musical Performance. *Ann. N. Y. Acad. Sci.* 423, 420–428.
- Shaffer, L. H. (1984b). Timing in solo and duet piano performances. *Q. J. Exp. Psychol. Sect. Hum. Exp. Psychol.* 36, 577.

- Shams, L., and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends Cogn. Sci.* 12, 411–417.
- Shea, C. H., Wulf, G., Park, J. H., and Gaunt, B. (2001). Effects of an auditory model on the learning of relative and absolute timing. *J. Mot. Behav.* 33, 127–138.
- Shea, J. B., and Morgan, R. L. (1979). Contextual Interference Effects on the Acquisition, Retention, and Transfer of a Motor Skill. *J. Exp. Psychol. [Hum. Learn.]* 5, 179–87.
- Sherrington, C. S. (1907). On the proprioceptive system, especially in its reflex aspect. *Brain* 29, 467–482.
- Sherwood, L. (1997). *Human physiology: from cells to systems*. Belmont, CA: Wadsworth Pub. Co.
- Shiller, D. M., Gracco, V. L., and Rvachew, S. (2010). Auditory-motor learning during speech production in 9-11-year-old children. *PLoS One* 5, e12975.
- Shiller, D. M., Sato, M., Gracco, V. L., and Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *J. Acoust. Soc. Am.* 125, 1103–1113.
- Shimoyama, I., Ninchoji, T., and Uemura, K. (1990). The Finger-Tapping Test: A Quantitative Analysis. *Arch Neurol* 47, 681–684.
- Shin, Y. K., Proctor, R. W., and Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychol. Bull.* 136, 943–974.
- Shi, Y. X., Tian, J. H., Yang, K. H., and Zhao, Y. (2011). Modified constraint-induced movement therapy versus traditional rehabilitation in patients with upper-extremity dysfunction after stroke: a systematic review and meta-analysis. *Arch. Phys. Med. Rehabil.* 92, 972–982.
- Siebner, H. R., Peller, M., Willoch, F., Minoshima, S., Boecker, H., Auer, C., Drzezga, A., Conrad, B., and Bartenstein, P. (2000). Lasting cortical activation after repetitive TMS of the motor cortex: a glucose metabolic study. *Neurology* 54, 956–963.
- Simoneau, G. G., Degner, R. M., Kramper, C. A., and Kittleson, K. H. (1997). Changes in Ankle Joint Proprioception Resulting From Strips of Athletic Tape Applied Over the Skin. *J. Athl. Train.* 32, 141–147.
- Sloboda (2000). Individual differences in music performance. *Trends Cogn. Sci.* 4, 397–403.
- Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Q. J. Exp. Psychol. Sect. 35*, 377–396.
- Smania, N., Paolucci, S., Tinazzi, M., Borghero, A., Manganotti, P., Fiaschi, A., Moretto, G., Bovi, P., and Gambarin, M. (2007). Active finger extension: a simple movement predicting recovery of arm function in patients with acute stroke. *Stroke J. Cereb. Circ.* 38, 1088–1090.
- Smith, S. M., Brown, H. O., Toman, J. E. P., and Goodman, L. S. (1947). The lack of cerebral effects of d-tubocurarine. *Anesthesiology* 8, 1–14.
- Sober, S. J., and Brainard, M. S. (2009). Adult birdsong is actively maintained by error correction. *Nat. Neurosci.* 12, 927–931.
- Sober, S. J., and Brainard, M. S. (2012). Vocal learning is constrained by the statistics of sensorimotor experience. *Proc. Natl. Acad. Sci.* 109, 21099–21103.
- Sokolov, A. N. (1972). *Inner Speech and Thought*. Basic Books.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psychol.* 43, 482–489.
- Stenneken, P., Prinz, W., Bosbach, S., and Aschersleben, G. (2006). Visual proprioception in the timing of movements: evidence from deafferentation. *Neuroreport* 17, 545–548.
- Stetson, C., Cui, X., Montague, P. R., and Eagleman, D. M. (2006). Motor-Sensory Recalibration Leads to an Illusory Reversal of Action and Sensation. *Neuron* 51, 651–659.
- Stevens, J. A., Fonlupt, P., Shiffrar, M., and Decety, J. (2000). New aspects of motion perception: selective neural encoding of apparent human movements. *Neuroreport* 11, 109–115.
- Stinear, C. M., Barber, P. A., Smale, P. R., Coxon, J. P., Fleming, M. K., and Byblow, W. D. (2007). Functional potential in chronic stroke patients depends on corticospinal tract integrity. *Brain J. Neurol.* 130,

170–180.

- Stöcker, C., Sebald, A., and Hoffmann, J. (2003). The influence of response--effect compatibility in a serial reaction time task. *Q. J. Exp. Psychol. A* 56, 685–703.
- Stuart, A., Frazier, C. L., Kalinowski, J., and Vos, P. W. (2008). The Effect of Frequency Altered Feedback on Stuttering Duration and Type. *J. Speech Lang. Hear. Res.* 51, 889–897.
- Stuart, A., Kalinowski, J., Rastatter, M. P., and Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *J. Acoust. Soc. Am.* 111, 2237–2241.
- Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., and Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain Cogn.* 82, 127–136.
- Sugano, Y., Keetels, M., and Vroomen, J. (2010). Adaptation to motor-visual and motor-auditory temporal lags transfer across modalities. *Exp. Brain Res. Exp. Hirnforsch. Expérimentation Cérébrale* 201, 393–399.
- Suteerawattananon, M., Morris, G. S., Etnyre, B. R., Jankovic, J., and Protas, E. J. (2004). Effects of visual and auditory cues on gait in individuals with Parkinson’s disease. *J. Neurol. Sci.* 219, 63–69.
- Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N., and Castiello, U. (2004). The Human Premotor Cortex Is “Mirror” Only for Biological Actions. *Curr. Biol.* 14, 117–120.
- Tan, S. S., Maul, T. H. B., and Mennie, N. R. (2013). Measuring the performance of visual to auditory information conversion. *PLoS One* 8, e63042.
- Taub, E. (2012). The behavior-analytic origins of constraint-induced movement therapy: an example of behavioral neurorehabilitation. *Behav. Anal. MABA* 35, 155–178.
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., and Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* 17, 273–281.
- Thaler, L., and Goodale, M. A. (2010). The Role of Online Visual Feedback for the Control of Target-Directed and Allocentric Hand Movements. *J. Neurophysiol.* 105, 846–859.
- Thaut, M. H., Kenyon, G. P., Hurt, C. P., McIntosh, G. C., and Hoemberg, V. (2002). Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients. *Neuropsychologia* 40, 1073–1081.
- Thaut, M. H., Leins, A. K., Rice, R. R., Argstatter, H., Kenyon, G. P., McIntosh, G. C., Bolay, H. V., and Fetter, M. (2007). Rhythmic auditory stimulation improves gait more than NDT/Bobath training in near-ambulatory patients early poststroke: a single-blind, randomized trial. *Neurorehabil. Neural Repair* 21, 455–459.
- Theidon, T., Mandrick, K., Foissac, M., Mottet, D., and Perrey, S. (2011). Degraded postural performance after muscle fatigue can be compensated by skin stimulation. *Gait Posture* 33, 686–689.
- Thomas, L. E. (2013). Spatial Working Memory Is Necessary for Actions to Guide Thought. *J. Exp. Psychol. Learn. Mem. Cogn.*, No Pagination Specified.
- Thompson, W. F., Schellenberg, E. G., and Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychol. Sci.* 12, 248–251.
- Timm, J., SanMiguel, I., Saupe, K., and Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neurosci.* 14, 2.
- Tkach, D., Reimer, J., and Hatsopoulos, N. G. (2007). Congruent Activity during Action and Action Observation in Motor Cortex. *J. Neurosci.* 27, 13241–13250.
- Todd, N. (1985). A model of expressive timing in tonal music. *Music Percept.*, 33–57.
- Tomasino, B., and Rumiati, R. I. (2013). At the Mercy of Strategies: The Role of Motor Representations in Language Understanding. *Front. Psychol.* 4. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3562995/> [Accessed August 26, 2013].
- Toni, I., de Lange, F. P., Noordzij, M. L., and Hagoort, P. (2008). Language beyond action. *J. Physiol.-Paris* 102, 71–79.
- Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably

- different? *Learn. Motiv.* 1, 129–140.
- Tsay, C.-J. (2013). Sight over sound in the judgment of music performance. *Proc. Natl. Acad. Sci.* Available at: <http://www.pnas.org/content/early/2013/08/16/1221454110> [Accessed August 26, 2013].
- Turk, A. E., and Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *J. Phon.* 35, 445–472.
- Uithol, S., van Rooij, I., Bekkering, H., and Haselager, P. (2011). Understanding motor resonance. *Soc. Neurosci.* 6, 388–397.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., and Rizzolatti, G. (2001). I Know What You Are Doing. *Neuron* 31, 155–165.
- Urgesi, C., Candidi, M., Ionta, S., and Aglioti, S. M. (2007). Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nat. Neurosci.* 10, 30–31.
- Urner, M., Sarri, M., Grahn, J., Manly, T., Rees, G., and Friston, K. (2013). The role of prestimulus activity in visual extinction. *Neuropsychologia*.
- Vahdat, S., Darainy, M., Milner, T. E., and Ostry, D. J. (2011). Functionally Specific Changes in Resting-State Sensorimotor Networks after Motor Learning. *J. Neurosci.* 31, 16907–16915.
- Vickhoff, B., Malmgren, H., Aström, R., Nyberg, G., Ekström, S.-R., Engwall, M., Snygg, J., Nilsson, M., and Jörnsten, R. (2013). Music structure determines heart rate variability of singers. *Front. Psychol.* 4, 334.
- Vidoni, E. D., and Boyd, L. A. (2009). Preserved motor learning after stroke is related to the degree of proprioceptive deficit. *Behav. Brain Funct. BBF* 5, 36.
- Vink, A. C., Birks, J., Bruinsma, M. S., and Scholten, R. J. (2003). “Music therapy for people with dementia,” in *Cochrane Database of Systematic Reviews*, eds. The Cochrane Collaboration and A. C. Vink (Chichester, UK: John Wiley & Sons, Ltd). Available at: <http://www2.cochrane.org/reviews/en/ab003477.html> [Accessed March 10, 2011].
- Vossel, S., Weiss, P. H., Eschenbeck, P., Saliger, J., Karbe, H., and Fink, G. R. (2012). The neural basis of anosognosia for spatial neglect after stroke. *Stroke J. Cereb. Circ.* 43, 1954–1956.
- Wagle Shukla, A., Moro, E., Gunraj, C., Lozano, A., Hodaie, M., Lang, A., and Chen, R. (2013). Long-term subthalamic nucleus stimulation improves sensorimotor integration and proprioception. *J. Neurol. Neurosurg. Psychiatry* 84, 1020–1028.
- Wagle-Shukla, A., Ni, Z., Gunraj, C. A., Bahl, N., and Chen, R. (2009). Effects of short interval intracortical inhibition and intracortical facilitation on short interval intracortical facilitation in human primary motor cortex. *J. Physiol.* 587, 5665–5678.
- Wagner, C. (1971). “The influence of the tempo of playing on the rhythmic structure studied at pianist’s playing scales,” in *Medicine and sport Biomechanics II.*, eds. J. Vredenburg and J. Wartenweiler (Basel: Karger), 129–132.
- Waldstein, R. S. (1990). Effects of postlingual deafness on speech production: implications for the role of auditory feedback. *J. Acoust. Soc. Am.* 88, 2099–2114.
- Ward, N. S., Brown, M. M., Thompson, A. J., and Frackowiak, R. S. J. (2003). Neural correlates of outcome after stroke: a cross-sectional fMRI study. *Brain J. Neurol.* 126, 1430–1448.
- Ward, N. S., Brown, M. M., Thompson, A. J., and Frackowiak, R. S. J. (2004). The influence of time after stroke on brain activations during a motor task. *Ann. Neurol.* 55, 829–834.
- Ward, N. S., and Cohen, L. G. (2004). Mechanisms underlying recovery of motor function after stroke. *Arch. Neurol.* 61, 1844–1848.
- Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J. S., Rosen, S., and Scott, S. K. (2006). Positive Emotions Preferentially Engage an Auditory–Motor “Mirror” System. *J. Neurosci.* 26, 13067–13075.
- Wassinger, C. A., Myers, J. B., Gatti, J. M., Conley, K. M., and Lephart, S. M. (2007). Proprioception and Throwing Accuracy in the Dominant Shoulder After Cryotherapy. *J. Athl. Train.* 42, 84–89.

- Watkins, K. E., Strafella, A. P., and Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41, 989–994.
- Weidler, B. J., and Abrams, R. A. (2013). Hand proximity—not arm posture—alters vision near the hands. *Atten. Percept. Psychophys.* 75, 650–653.
- Wei, K., and Körding, K. (2009). Relevance of Error: What Drives Motor Adaptation? *J. Neurophysiol.* 101, 655–664.
- Weinrich, M., and Wise, S. P. (1982). The premotor cortex of the monkey. *J. Neurosci. Off. J. Soc. Neurosci.* 2, 1329–1345.
- Weisberg, J., Turennout, M. van, and Martin, A. (2007). A Neural System for Learning about Object Function. *Cereb. Cortex* 17, 513–521.
- Weiskrantz, L., Elliott, J., and Darlington, C. (1971). Preliminary Observations on Tickling Oneself. *Nature* 230, 598–599.
- Werhahn, K. J., Mortensen, J., Van Boven, R. W., Zeuner, K. E., and Cohen, L. G. (2002). Enhanced tactile spatial acuity and cortical processing during acute hand deafferentation. *Nat. Neurosci.* 5, 936–938.
- Werker, J. F., and Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends Cogn. Sci.* 9, 519–527.
- Wiesendanger, M. (1981). “Organization of Secondary Motor Areas of Cerebral Cortex,” in *Comprehensive Physiology* (John Wiley & Sons, Inc.). Available at: <http://onlinelibrary.wiley.com/doi/10.1002/cphy.cp010224/abstract> [Accessed May 30, 2013].
- Wilson, S. M., and Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *NeuroImage* 33, 316–325.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., and Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7, 701–702.
- Wiltermuth, S. S. (2012). Synchrony and destructive obedience. *Soc. Influ.* 7, 78–89.
- Wiltermuth, S. S., and Heath, C. (2009). Synchrony and Cooperation. *Psychol. Sci.* 20, 1–5.
- Winter, J. A., Allen, T. J., and Proske, U. (2005). Muscle Spindle Signals Combine with the Sense of Effort to Indicate Limb Position. *J. Physiol.* 568, 1035–1046.
- Wise, S. P., and Mauritz, K. H. (1985). Set-related neuronal activity in the premotor cortex of rhesus monkeys: effects of changes in motor set. *Proc. R. Soc. Lond. Ser. B Contain. Pap. Biol. Character R. Soc. Gt. Br.* 223, 331–354.
- Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nat. Rev. Neurosci.* 12, 739–751.
- Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882.
- Wolpert, D. M., Miall, R. C., and Kawato, M. (1998). Internal models in the cerebellum. *Trends Cogn. Sci.* 2, 338–347.
- Wong, J. D., Kistemaker, D. A., Chin, A., and Gribble, P. L. (2012). Can proprioceptive training improve motor learning? *J. Neurophysiol.* 108, 3313–3321.
- Wu, D., Li, C.-Y., and Yao, D.-Z. (2009). Scale-free music of the brain. *PLoS One* 4, e5915.
- Yates, A. J. (1963). Delayed auditory feedback. *Psychol. Bull.* 60, 213–232.
- Zatorre, R. J. (2007). There’s more to auditory cortex than meets the ear. *Hear. Res.* 229, 24–30.
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci* 8, 547–558.
- Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., and Evans, A. C. (1996). Hearing in the Mind’s Ear: A PET Investigation of Musical Imagery and Perception. *J. Cogn. Neurosci.* 8, 29–46.
- Ziessler, M. (1998). Response–effect learning as a major component of implicit serial learning. *J. Exp. Psychol.*

Sounds on time - References and indices

*Learn. Mem. Cogn.* 24, 962–978.

## Alphabetical Index

### A

Active sensing.....	34
Activities of daily living.....	78
Afferent.....	23
Agrammatic aphasia.....	53
Associative Sequence Learning.....	49
Audiospinal facilitation.....	76

### B

Basal ganglia.....	24
Basal nuclei.....	24
Behaviourism.....	31
Bimanual advantage.....	78
Broca's aphasia.....	53

### C

Cerebellum.....	24
Cerebrovascular incident.....	164
Cognitive grammar.....	54
Common coding.....	71
Conduction aphasia.....	80
Constraint-induced movement therapy.....	167
Contralateral inhibition.....	165
Convergence.....	56
Corollary discharge.....	37
Corticospinal tract integrity.....	165
Counter-mirror.....	48
Coupling.....	67

### D

Differential outcome effect.....	72
Direct Matching Hypothesis.....	47
Distal muscles.....	25

### E

Ecological musical stimuli.....	253
Efference copy.....	36
Efferent.....	23
Electromyography.....	23
Embodiment.....	31
EMG.....	23
Error signal.....	26
Expressive aphasia.....	53
Extinction.....	166

### F

Feedback.....	26, 67
Feedforward.....	26
Fluent aphasia.....	53

### G

Groove.....	61
Guidance hypothesis.....	74

### H

H-reflex.....	75
Haemorrhage.....	164
Hemiparesis.....	164
Hoffman's reflex.....	75

### I

Ideomotor principle.....	70
--------------------------	----

Ideomotor theory.....	70	Music.....	58
Inner speech.....	31	Music-supported therapy.....	172
Internal forward models.....	27	Musician's dystonia.....	254
Internal simulation.....	62	<b>N</b>	
Intra-cortical facilitation.....	63	Neglect.....	166
Iowa piano camera.....	250	Neuronal reorganisation.....	166
Ischaemia.....	164	Nociception.....	170
<b>J</b>		Non-fluent aphasia.....	53
Joint position sense.....	168	<b>P</b>	
<b>K</b>		Parkinson's disease.....	78
Keystroke velocity.....	250	Perception-for-Action-Control Theory.....	51
Kinesthesia.....	168	Perceptual fading.....	35
<b>L</b>		Perceptual hypothesis.....	251
Learned non-use.....	167	Peri-lesional areas.....	165
Lexical frequency effect.....	52	Phonemes.....	50
Local structure computation.....	54	Piano roll.....	250
<b>M</b>		Premotor cortex.....	25
Mental motor imagery.....	62	Primary motor cortex.....	24
Microsaccades.....	35	Proprioception.....	168
MIDI.....	250	Proximal muscles.....	25
Mirror network.....	43	<b>R</b>	
Mirror neurons.....	39	Receptive aphasia.....	53
Mismatch-negativity.....	60	Reference-frame transformation.....	73
Motor learning.....	73	Repositioning sense.....	169
Motor resonance.....	46	Retention test.....	73
Motor theory of speech perception.....	50	<b>S</b>	
Motor unit.....	23	Sensorimotor contingencies.....	32
Motor-evoked potentials.....	25	Sensorimotor integration.....	28
Motor-induced suppression.....	36	Sensorimotor synchronisation.....	77
Muscle fibres.....	23	Sensory attenuation.....	36
Muscle spindles.....	23	Sensory feedback.....	66

Sensory stimulation.....67p.	<b>T</b>
Serial reaction time.....59	Temporal binding.....38
Serial Reaction Time.....89	Theory of event coding.....70
Set-related response.....42	Thumb Localisation Test.....169
Small-world.....22	Trans-cranial magnetic stimulation.....25
Somatotopic.....52	Transfer.....329
Sonification.....67, 82	Troxler effect.....35
State Feedback Control.....27	Two-phase model of action control.....71
Stretch receptors.....23	<b>U</b>
Stroke.....164	Unevenness.....253
Subvocal speech.....31	<b>W</b>
Supplementary motor area.....25	Wernicke's aphasia.....53

## Illustration Index

Illustration 1: Feedback-based learning.....	27
Illustration 2: Element vs. sequence integration.....	34
Illustration 3: Explanation of the difference between mirror networks and mirror neurons.....	44
Illustration 4: Sensory feedback vs. sensory stimulation.....	68
Illustration 5: Contrasting regularity and frequency (speed). Sequence production timing can, in principle, vary independently in speed and in regularity.....	89
Illustration 6: Summary of findings in typical Serial Reaction Time studies.....	90
Illustration 7: Explanation of the perceptual hypothesis. By hypothesis, a systematic perceptual bias exists that makes certain intervals sound longer (or shorter) than others. In our example, the last interval sounds shorter than the previous intervals. If the musician's task is to produce as regular a scale as possible, she will play the final interval longer so that all intervals sound equally long.....	252
Illustration 8: Visualisation of the scale playing movement. Photos are ordered from left to right chronologically.....	253
Illustration 9: Interval-based Scale Analysis.....	256
Illustration 10: Interval-based vs. note-based timing.....	257
Illustration 11: Artistic impression of the pianistic fingerprint. Drawing by Vera Scepanovic, based on an idea by Mariana Fulgueiras Montoro.....	281
Illustration 12: Summary of the irregularity-instability model of timing in scale playing.....	319



## **VII. Appendix: Parkinson's Disease patients**

**VII.1. Effects of dopaminergic and subthalamic stimulation  
on musical performance.**

---













## **VIII.Appendix: Fine-tuning the timing analysis in experts**

### **VIII.1. The influence of chronotype on making music: Circadian fluctuations in pianists' fine motor skills**

---

This study reveals the influence of circadian fluctuations in motor timing in pianists, and further establishes the irregularity-instability analysis method for piano scale playing.



# The influence of chronotype on making music: circadian fluctuations in pianists' fine motor skills

Floris T. Van Vugt<sup>1,2</sup>, Katharina Treutler<sup>1</sup>, Eckart Altenmüller<sup>1</sup> and Hans-Christian Jabusch<sup>3\*</sup>

<sup>1</sup> Institute of Music Physiology and Musicians' Medicine, University of Music, Drama, and Media, Hanover, Germany

<sup>2</sup> Lyon Neuroscience Research Center, CNRS-UMR 5292, INSERM U1028, University Lyon-1, Lyon, France

<sup>3</sup> Institute of Musicians' Medicine, University of Music Carl Maria von Weber, Dresden, Germany

## Edited by:

Shinichi Furuya, Hanover University of Music, Drama and Media, Germany

## Reviewed by:

Antoni Rodriguez-Fornells, University of Barcelona, Spain  
Martha Merrow, LMU, Germany

## \*Correspondence:

Hans-Christian Jabusch, Institute of Musicians' Medicine, University of Music Carl Maria von Weber, Dresden, Leubnitzer Strasse 17b, 01069 Dresden, Germany  
e-mail: jabusch@hfmdd.de

Making music on a professional level requires a maximum of sensorimotor precision. Chronotype-dependent fluctuations of sensorimotor precision in the course of the day may prove a challenge for musicians because public performances or recordings are usually scheduled at fixed times of the day. We investigated pianists' sensorimotor timing precision in a scale playing task performed in the morning and in the evening. Participants' chronotype was established through the Munich Chrono-Type Questionnaire, where mid-sleep time served as a marker for the individual chronotypes. Twenty-one piano students were included in the study. Timing precision was decomposed into consistent within-trial variability (irregularity) and residual, between-trial variability (instability). The timing patterns of late chronotype pianists were more stable in the evening than in the morning, whereas early chronotype pianists did not show a difference between the two recording timepoints. In sum, the present results indicate that even highly complex sensorimotor tasks such as music playing are affected by interactions between chronotype and the time of day. Thus, even long-term, massed practice of these expert musicians has not been able to wash out circadian fluctuations in performance.

**Keywords:** musician, piano, scale playing, practice, sensorimotor performance, chronobiology, chronotype, circadian fluctuation

## INTRODUCTION

Periodic processes within biological organisms occur at various timescales, such as seconds (e.g., heartbeat or respiration), days (e.g., the sleep cycle), weeks (e.g., the circaseptan rhythms), or months (e.g., menstruation cycle). Those processes that operate on a roughly 24-h cycle are governed by the circadian clock, which provides a temporal structure that modulates biological functions to match the daily cycle with the environment (Roenneberg and Merrow, 2003; Merrow et al., 2005). The circadian clock is entrained to the 24 h in a day by so-called zeitgebers such as the light-dark cycle (Roenneberg and Merrow, 2003) but also other non-oscillating environmental factors (Roenneberg et al., 2003a).

Circadian fluctuations are evident in various physiological functions of the human organism: clinical chemical parameters and endocrinological parameters such as concentration of hemoglobin, potassium, iron, adrenaline, noradrenaline, cortisol, and other hormones in blood and serum (Wisser and Breuer, 1981), body temperature (Aschoff, 1955) as well as cognitive functions such as reaction time to sensory cues (Kleitman et al., 1938) and memory tasks of various complexity (Van Eekelen and Kerkhof, 2003). In particular, circadian fluctuations also occur within the sensorimotor system. Circadian maxima and minima have, for example, been reported in grip strength (Atkinson et al., 1993), elbow flexion torque (Gauthier et al., 1996), back and leg strength (Coldwells et al., 1994), finger-tapping (Dosseville et al., 2002), tracking tasks (Van Eekelen and Kerkhof, 2003),

and manual dexterity (Monk and Kupfer, 2000). However, to our knowledge, no studies exist that investigate variations in musical performance according to the circadian cycle.

Investigation of circadian rhythms has an important caveat: considerable inter-individual differences exist in the circadian clock as well as in its entrainment. These differences result in individual preferences in the timing of sleep and wake commonly referred to as chronotypes. Different chronotypes have been described such as "larks" (early sleepers) and "owls" (late sleepers) (Roenneberg et al., 2003a,b). Questionnaires have been developed to establish the individual chronotype either by assigning people to various categories such as morning, evening and intermediate type categories (Horne and Ostberg, 1976). Alternatively, researchers have calculated the midpoint between sleep onset and wake up, referred to as the mid-sleep timepoint, as the phase reference point for the sleep cycle (Benoit et al., 1981; Roenneberg et al., 2003b).

Indeed, these inter-individual differences in the circadian clock may explain differences in circadian fluctuations. This has been reported for cognitive performance in a range of memory tasks (Petros et al., 1990; May et al., 1993; Hasher et al., 1999; Intons-Peterson et al., 1999; West et al., 2002), for the alerting component of attention (Matchock and Mordkoff, 2009) as well as for the sensorimotor system, revealing chronotype-induced changes in maximum voluntary muscle contraction and excitability of the motor cortex (Tamm et al., 2009), furthermore in influence of bright light on physical performance (Kantermann et al., 2012).

Activity patterns of the neural networks during the day have also been reported to be chronotype-dependent (Peres et al., 2011). As a consequence of the chronotype-dependent properties of circadian performance fluctuations, recent studies on circadian rhythms in the performance of complex tasks were based on study designs that controlled for the chronotype. For example, such chronotype-controlled circadian studies were used to identify dual-task costs (Jasper et al., 2010) or soccer-related performance (Reilly et al., 2007).

Performing music is a complex task executed on different times of the day. Performing music at a professional level is regarded as one of the most complex tasks in human life (Münste et al., 2002). Playing in public requires the highest possible level of performance independent of the time of the day. In matinee concerts, expert musicians have to play in the morning whereas in evening concerts, they have to perform at night.

To our knowledge, to date no systematic investigation exists into performance fluctuations across the daily cycle in musicians and their potential association with the chronotype. The present study investigates the performance quality in a demanding music-related sensorimotor task in professional pianists and its association with the time of the day and the chronotype. In order to objectively quantify performance quality, pianists were measured playing musical scales (Jabusch et al., 2004). Scales are never played perfectly evenly, therefore we divided timing deviations in systematic deviations (*irregularity*) and trial-to-trial variability (*instability*) (van Vugt et al., 2012). The temporal deviations that are consistent within trials (irregularity) were previously found to be inaudible and mostly determined by neuromuscular constraints (van Vugt et al., 2013). Therefore, we hypothesize that this part of the variability does not change across the daily cycle. However, instability reflects trial-to-trial variability away from these consistent deviations and is therefore a likely candidate for circadian fluctuations. We therefore hypothesized that early sleepers would be more stable, but not more regular, in the morning than in the evening. Similarly, late sleepers are expected to be more stable, but not more regular, in the evening than in the morning.

## METHODS

### PARTICIPANTS

22 piano students (8 females) were recruited from the student pool at the University of Music, Drama and Media in Hanover. Participants were 22.5 ( $SD = 2.9$ ) years old. All were right-handed according to the Edinburgh Handedness Inventory except for two left-handed participants (laterality quotient:  $M = 90.7$ ,  $SD = 11.6$  in right-handed participants;  $-81.8$  and  $-100$  for the two left-handed participants). Two participants were in the “piano performance” study path and one was studying in the so-called solo class (postgraduate studies); all others were in the “music education” study path which has lower admission standards. None of the participants reported any neurological condition. Participants were not selected using chronotype criteria. Rather than focusing on very early or very late sleepers (larks or owls), we decided to investigate circadian fluctuations in a representative sample of music students.

### PROCEDURE

Participants came to the lab on two different days and were recorded playing scales. One measurement happened at 8 AM (henceforth referred to as AM recording) and the other at 8 PM (henceforth referred to as PM recording). Participants avoided to play piano prior to the measurement on both study days except a warm-up of five minutes immediately before the recording. The order of the AM and PM measurements was counterbalanced. Participants played on a MP 9000 MIDI stage piano (Kawai, Krefeld, Germany). The keyboard’s digital music interface (MIDI out) signal was captured on a PC using a commercially available sequencer software (Musicator Win, version 2.12; Music Interactive Technology, Bergen, Norway).

Participants were requested to play two-octave C-major scales beginning with the C (131 Hz) one octave below the middle C and ending with the C (523 Hz) one octave higher than the middle C. Ascending and descending scales were interleaved. The instruction to the participants was to play as evenly as possible and in a legato style at mezzo-forte loudness. In order to maximally challenge the motor system, participants were required to play fast. This was achieved through presenting a metronome beat at 160 BPM and instructing the participants to play at 4 notes per metronome beat, resulting in 10.7 keystrokes per second. Participants performed roughly 15 scales with the right hand and with the left hand using the conventional fingering (123123412312345 and reverse, where the numbers indicate the fingers from the thumb, 1, to the little finger, 5).

### QUESTIONNAIRES: CHRONOTYPE AND PRACTICE HABITS

Sleep habits were assessed applying the Munich Chrono-Type Questionnaire (MCTQ) (Roenneberg et al., 2003b), which participants filled out after whichever of the two recording sessions was last. This questionnaire is a validated tool to identify the chronotype based on self-reported individual sleep times, considering work and free days separately. The mid-point between sleep onset and waking up serves as sleep phase reference point. The sleep phase reference point is identified following an established protocol reported by Roenneberg et al. (2004): The so-called mid-sleep time-point is identified for the sleep in nights before work days and referred to as MSW. Additionally, the mid-sleep time point for nights before free days (MSF) is identified. To adjust the mid-sleep time point for the fact that individuals typically accumulate sleep debt on work days and compensate for this on free days, an adjusted mid-sleep (MSFsc) is calculated as follows. We first calculate the average daily sleep duration or need (ASD) as follows:  $ASD = (X \times SDW + (7 - X) \times SDF)/7$  where X is the number of work days per week, SDF is sleep duration on free days and SDW is sleep duration on work days. Then the adjusted mid-sleep time point is calculated as follows:  $MSFsc = MSF - 0.5 \times (SDF - ASD)$ .

Additionally, a researcher-developed questionnaire focused on practice history and temporal practice habits at present. In analogy to mid-sleep, the individual “mid-practice time” (MPT) was identified. Participants reported their amounts of piano practice for each of eight 3-h time windows throughout

the 24-h day. Time-windows were weighted according to these respective practice amounts. This enabled us to identify the MPT-point as the mid-point in time between onset and end of daily practice taking into account rests in between practice sessions.

### SCALE PLAYING ANALYSIS

First, we calculated the *unevenness* measure that has traditionally been computed to assess temporal precision in scale playing (Wagner, 1971; Jabusch et al., 2004). We proceeded to calculate the note onsets for each correct scale run and the standard deviation of the inter-onset intervals (in ms). We calculated the median of these for each combination of hand, playing direction (inward or outward; inward being defined as radial playing direction, outward as ulnar playing direction) and recording within each participant. As a secondary measure, we recorded the downward velocity of the keystrokes (in MIDI units). This parameter indirectly influences loudness: the faster the keystroke, the louder the sound. For each scale run, we calculated the SD of the keystroke velocities and then pooled these by taking the median for each pianist and condition.

Secondly, we established note-by-note timing in scale playing according to a protocol published previously (van Vugt et al., 2012). First, we isolated correctly performed scale runs, discarding those containing errors or additional notes. We then converted the note values to their rank in the C major scale (i.e., C has rank 0, D has rank 1, E has rank 2, etc., up to C'' with rank 14) and performed a least-square straight line fit to this set of pairs of rank and timing. This allowed us to compute for each note the expected onset time (according to this fit) and then the deviation of the timing of the actually measured onset (in ms). We performed this fit for all scale runs and then pooled the results by hand (left or right), playing direction (inward or outward) and note, calculating the median lateness (*irregularity*; in ms) as well as inter-quartile range (*instability*; in ms) for that condition. Irregularity represents the amount of deviation away from regularity present within trials, capturing the fact that some notes are consistently late or consistently early within trials. Independent of this, the lateness of notes can vary between trials. Instability captures this by quantifying the variability in note timing across several trials. **Figure 1** illustrates this procedure in two example pianists. It shows, firstly, that the pianists have highly individual timing traces (van Vugt et al., 2013). Secondly, in this example, the timing traces are more consistent (stable) for the early chronotype pianist's AM recording compared to the PM recording. The opposite is true for the late chronotype pianist. It is this observation that the following analysis captures. For more details on this procedure, the reader is invited to consult van Vugt et al. (2012). In sum, we segmented for each pianist the timing variability into variability due to deviations that are present across trials (irregularity) and variability between trials (instability).

Our ANOVAs were Type-II and we report the generalized effect size  $\eta_G^2$  (Bakeman, 2005). Shapiro-Wilk normality test was used to verify normality of the data and Mauchly's test to detect sphericity violations, which were never significant for the data reported in this paper.

## RESULTS

### QUESTIONNAIRES: PRACTICE HISTORY, CURRENT PRACTICE HABITS AND CHRONOTYPE

One participant had to be eliminated due to technical reasons. The remaining 21 participants had started to play piano at the age of 5.5 ( $SD = 2.0$ ) years and had played the instrument for a total of 17.0 ( $SD = 3.1$ ) years, accumulating a total of 15.6 ( $SD = 7.5$ ) thousand practice hours. The current daily practice duration was 4.0 ( $SD = 1.9$ ) h. The median of the MPTs was 16.8 h local time, i.e., 4:48 PM (range, 13–20 h). The median number of work days per week was 7 (range, 4–7).

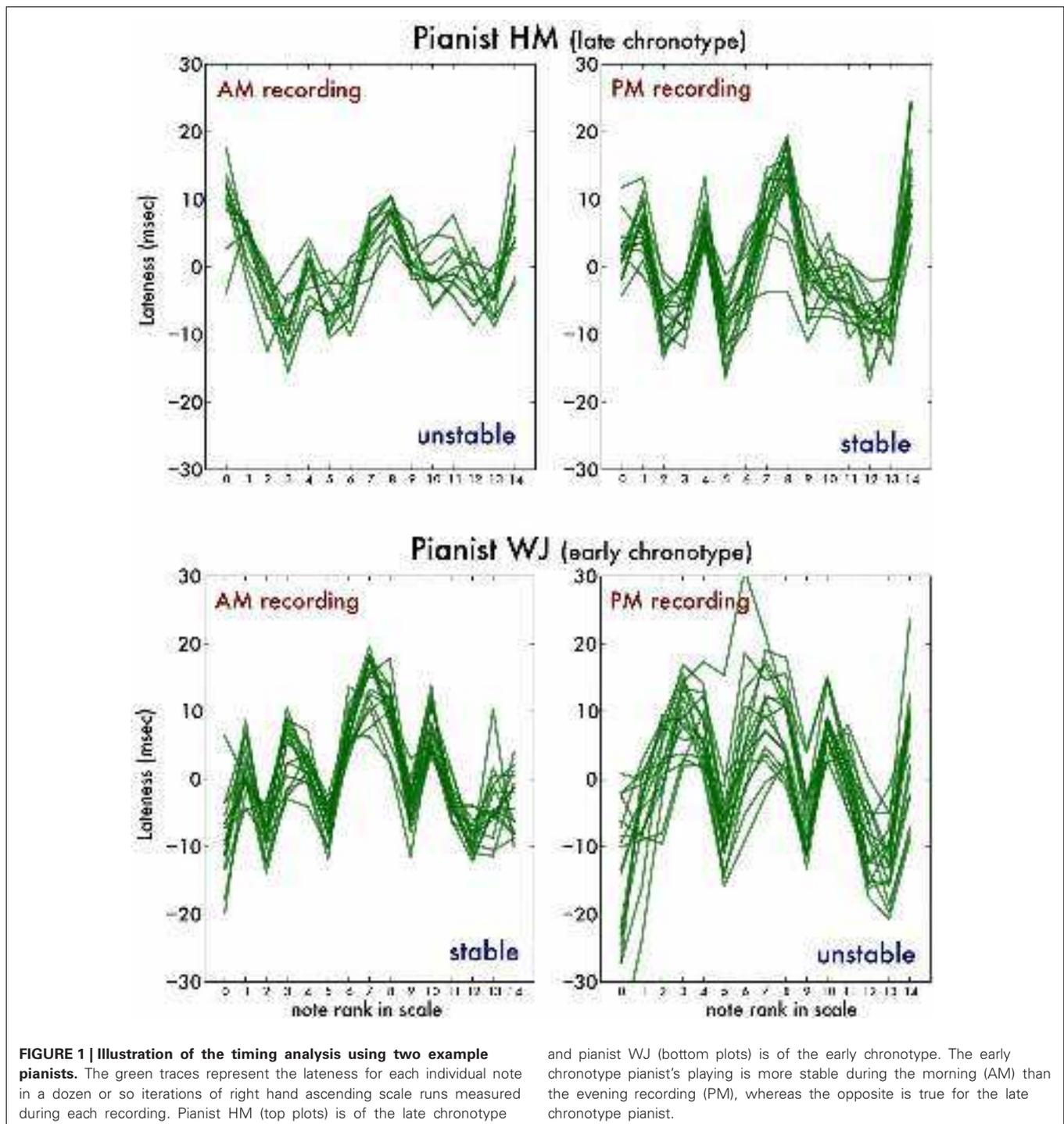
The MCTQ revealed an average sleep duration (total sleep duration on all work days and free days per week, divided by 7 days) of 7.7 ( $SD = 0.65$ ) h. The corrected mid-sleep time-point on free days (MSFsc) was used as a proxy for the participant chronotype. The median of MSFsc was 5.0 h past midnight, local time (5:00 AM) (range, 4.1–6.2 h). A correlation was seen between participant's mid-sleep time point on work days (MSW) and MPT [Pearson  $r_{(19)} = 0.65$ ,  $p = 0.001$ ] indicating that pianists with an early sleep phase during workday nights practiced early during the day and vice versa. There was no correlation between MPT and mid-sleep time point on free days (MSF) [Pearson  $r_{(19)} = 0.26$ ,  $p = 0.25$ ], nor the corrected mid-sleep time on free days (MSFsc) [Pearson  $r_{(19)} = 0.33$ ,  $p = 0.14$ ].

For further analysis, a median-split procedure was carried out to classify participants as either an earlier chronotype (MSFsc < 5.0) or a later chronotype subgroup (MSFsc  $\geq$  5.0).

To check that our median split division in early and late chronotypes was legitimate, we verified that there were no differences in gender, age, age of commencement of piano training, number of years of piano training, nor in accumulated practice hours at the piano between the two groups (**Table 1**). The two groups clearly differed in sleeping behavior by definition of the group division.

### SCALE PLAYING: EXTRACTING CORRECT SCALES

We recorded a total of 463 note onsets ( $SD = 53$ ) for each pianist, hand, and recording (AM or PM). More note onset material was recorded for the left ( $M = 482$ ,  $SD = 58$  onsets) than for the right hand ( $M = 446$ ,  $SD = 40$  onsets) [ $F_{(1, 19)} = 19.76$ ,  $p = 0.003$ ,  $\eta_G^2 = 0.06$ ]. From these, we extracted the correct inward and outward scales, discarding onsets that were part of incomplete scales or scales with errors (786 keystrokes discarded, amounting to 1.9% of the recorded material). We found 15.0 ( $SD = 2.50$ ) correctly produced scales for each participant, hand, playing direction (inward or outward) and recording. Participants' left hand playing contained more correct scales ( $M = 15.6$ ,  $SD = 1.75$ ) than the right hand ( $M = 14.6$ ,  $SD = 1.28$ ) [ $F_{(1, 19)} = 15.89$ ,  $p < 0.001$ ,  $\eta_G^2 = 0.05$ ]. There was no difference in number of correct scales according to the inward or outward playing direction [ $F_{(1, 19)} = 1.11$ ,  $p = 0.3$ ]. There was a trend for there to be more correctly played scales in the PM recording ( $M = 15.7$ ,  $SD = 2.16$  scales) than in the AM recording ( $M = 14.5$ ,  $SD = 2.30$  scales) [ $F_{(1, 19)} = 3.06$ ,  $p = 0.10$ ]. Importantly, there was no main effect of chronotype [ $F_{(1, 19)} = 0.01$ ,  $p = 0.9$ ], revealing that both chronotype subgroups played an equal number of correct scales. There was an interaction between playing direction



and time-point of recording [ $F_{(1, 19)} = 6.70, p = 0.02, \eta_G^2 = 0.002$ ], revealing that for the PM recording there were more correct inward than correct outward scales recorded, whereas for the AM recording the opposite was true. None of the other two-way interactions was significant.

#### SCALE PLAYING: UNEVENNESS

First we calculated the traditional measure of *unevenness*, defined as the standard deviation of the intervals between the onsets of

subsequent keystrokes. There was no main effect of playing direction (inward vs. outward) [ $F_{(1, 19)} = 0.06, p = 0.4$ ] or recording time-point [ $F_{(1, 19)} = 2.79, p = 0.11$ ]. However, there was a main effect of hand [ $F_{(1, 19)} = 26.35, p < 0.001, \eta_G^2 = 0.10$ ], indicating that the right hand played more evenly ( $M = 10.4, SD = 2.12$  ms) than the left hand ( $M = 12.3, SD = 2.64$  ms). There was no main effect of chronotype [ $F_{(1, 19)} = 0.03, p = 0.87$ ]. There was a trend for an interaction between hand and direction [ $F_{(1, 19)} = 4.14, p = 0.06, \eta_G^2 = 0.01$ ], indicating that for the

left hand, outward scales were more uneven than inward scales, whereas for the right hand the opposite was true. None of the two-way interactions were significant [all  $F_{(1, 19)} < 2.18, p > 0.15$ ], nor any of the three-way interactions [all  $F_{(1, 19)} < 1.65, p > 0.2$ ] or the four-way interaction [ $F_{(1, 19)} = 1.89, p = 0.18$ ].

**SCALE PLAYING: IRREGULARITY AND INSTABILITY**

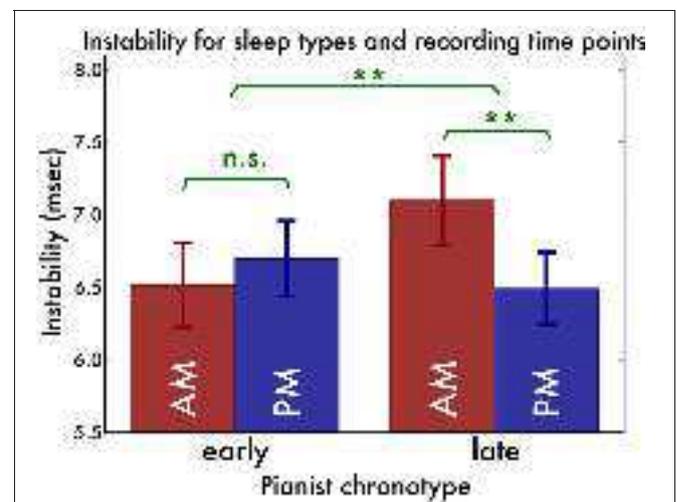
We proceeded to calculate the irregularity-instability analysis, which decomposes the variability in scale playing variability consistent timing deviations (*irregularity*) and trial-to-trial variability (*instability*) (van Vugt et al., 2012).

First, we report the results of the irregularity analysis. There was no main effect of pianist chronotype [ $F_{(1, 19)} = 0.02, p = 0.87$ ] or playing direction [ $F_{(1, 19)} = 1.18, p = 0.29$ ]. We found a trend for a main effect of recording time-point [ $F_{(1, 19)} = 3.31, p = 0.08$ ], showing that playing tended to be slightly more regular during the PM recordings ( $M = 6.14, SD = 1.41$  ms) than during the AM recordings ( $M = 6.54, SD = 1.59$  ms). Similarly, a trend for a main effect of hand [ $F_{(1, 19)} = 3.71, p = 0.07$ ] showed that right hand playing tended to be more regular ( $M = 6.08, SD = 1.53$  ms) than the left hand ( $M = 6.60, SD = 1.54$  ms). There was an interaction between hand and playing direction [ $F_{(1, 19)} = 5.81, p = 0.03, \eta_G^2 = 0.03$ ], indicating that the right hand outward scales were more regular than the inward scales, whereas for the left hand the opposite was true. Importantly, there was no interaction between pianist chronotype and recording time point [ $F_{(1, 19)} = 2.89, p = 0.11$ ]. None of the other two-, three- or four-way interactions was significant [all  $F_{(1, 19)} < 1.08, p > 0.3$ ].

The instability revealed a different picture. A main effect of hand [ $F_{(1, 19)} = 22.47, p < 0.001, \eta_G^2 = 0.14$ ] revealed that the left hand played more unstable ( $M = 7.16, SD = 1.11$  ms) than the right hand ( $M = 6.25, SD = 0.76$  ms). A main effect of direction [ $F_{(1, 19)} = 21.78, p < 0.001, \eta_G^2 = 0.03$ ] indicated that timing in outward scales was more stable ( $M = 6.50, SD = 0.79$  ms) than inward scales ( $M = 6.91, SD = 0.95$  ms). There was no main effect of pianist chronotype [ $F_{(1, 19)} = 0.25, p = 0.62$ ]. There was an interaction between scale direction and recording time-point [ $F_{(1, 19)} = 4.52, p = 0.05, \eta_G^2 = 0.01$ ], which indicated that

inward scales were more stable at the PM recording, whereas outward scales were equally stable during both recordings. Now, crucially, we found an interaction effect between pianist chronotype and recording time-point [ $F_{(1, 19)} = 10.20, p = 0.004, \eta_G^2 = 0.03$ ], which indicated that pianists of the late chronotype played more stably in the PM recording than in the AM recording [ $F_{(1, 10)} = 15.92, p = 0.003, \eta_G^2 = 0.06$ ] whereas pianists of the early chronotype played equally stable in both recordings [ $F_{(1, 9)} = 0.85, p = 0.38$ ] (Figure 2).

In order to assess how sensitive this result is to the intermediate chronotypes (those close to the median split), we performed a control analysis. In this analysis, we included only the 7 earliest chronotype pianists and the 7 latest chronotype pianists. With instability as a dependent variable, the same interaction between chronotype and recording time point was



**FIGURE 2 |** Playing instability (trial-to-trial variability) for the early and late chronotypes and the morning (red bars) and evening (blue bars) recordings. We found that early sleepers' playing was equally stable in both recordings, whereas late sleepers' playing was more stable in the evening than in the morning. \* $p < 0.05$ , \*\* $p < 0.01$ .

**Table 1 |** Characterization of the late- and early chronotype groups.

	Early chronotype	Late chronotype	Statistical comparison
Gender (female/male)	4/6	4/7	Fisher exact test $p = 1.0$
Age (years)	22.2 (3.65)	22.7 (2.28)	$t_{(14.8)} = -0.39, p = 0.70$
Handedness (Edinburgh laterality quotient %)	56.8 (78.6)	88.5 (11.53)	$t_{(9.4)} = -1.26, p = 0.24$
Age of commencement of piano training (years)	5.7 (1.87)	5.4 (2.25)	$t_{(18.9)} = 0.37, p = 0.71$
Amount of piano training (years)	16.5 (3.71)	17.3 (2.54)	$t_{(15.7)} = -0.62, p = 0.55$
Accumulated practice time (×1000 h)	16.3 (9.07)	15.0 (6.00)	$t_{(15.4)} = 0.37, p = 0.71$
Mid-sleep time before work days (MSW, hours after midnight)	4.4 (0.36)	4.9 (0.63)	$t_{(16.1)} = -2.23, p = 0.04$
Mid-sleep time before free days (MSF, hours after midnight)	5.0 (0.41)	5.6 (0.44)	$t_{(18.9)} = -3.09, p = 0.006$
Corrected mid-sleep time on free days (MSFsc, local time in hours after midnight)	4.6 (0.22)	5.3 (0.47)	$t_{(14.4)} = -4.68, p = 0.0003$
Mid-practice time (MPT, hours)	15.9 (2.31)	16.8 (1.35)	$t_{(14.2)} = -1.03, p = 0.32$
Average sleep duration (hours)	7.8 (0.74)	7.7 (0.41)	$t_{(13.7)} = 0.37, p = 0.72$

Values are reported as mean (SD) unless otherwise indicated.

found [ $F_{(2, 12)} = 7.17$ ,  $p = 0.02$ ,  $\eta_G^2 = 0.05$ ]. Again, no such interaction was present with irregularity as dependent variable [ $F_{(2, 12)} = 2.21$ ,  $p = 0.16$ ].

In order to eliminate problems due to the small sample size of our two groups, we performed the following analysis of covariance (ANCOVA) on our entire participant pool. First, we calculated the difference in irregularity and instability between the morning and evening recording for each pianist, hand and direction. We then performed a ANCOVA with the irregularity difference as outcome variable and the pianist chronotype as independent variable, allowing for variable slope and intercept according to hand and direction. The model did not reach significance [ $F_{(5, 78)} = 1.69$ ,  $p = 0.15$ ], further supporting the notion that irregularity differences between morning and evening recordings are not systematically influenced by pianist chronotype. We then performed the same regression with instability as a dependent variable and found a significant model [ $F_{(5, 78)} = 2.59$ ,  $p = 0.03$ ]. The mid-sleep time point on free days (MSFsc) was a significant predictor of the morning-evening instability difference [ $t = -2.24$ ,  $p = 0.03$ ] (Figure 3). The slope and intercept of the regression line were not different according to hand [both  $t < 0.13$ ,  $p > 0.89$ ] or playing direction [both  $t < 0.91$ ,  $p > 0.37$ ]. This regression corresponds to a negative correlation [ $r_{(19)} = -0.52$ ,  $p = 0.01$ ]. The same analysis with unevenness as outcome variable did not yield a significant model [ $F_{(5, 78)} = 0.78$ ,  $p = 0.56$ ] nor a correlation [ $r_{(19)} = -0.22$ ,  $p = 0.33$ ].

### SCALE PLAYING: KEYSTROKE VELOCITY

Thus far, we have restricted our attention to timing precision of the note onsets. Are other musical parameters affected? For each scale run, we took the average keystroke velocity and then took the

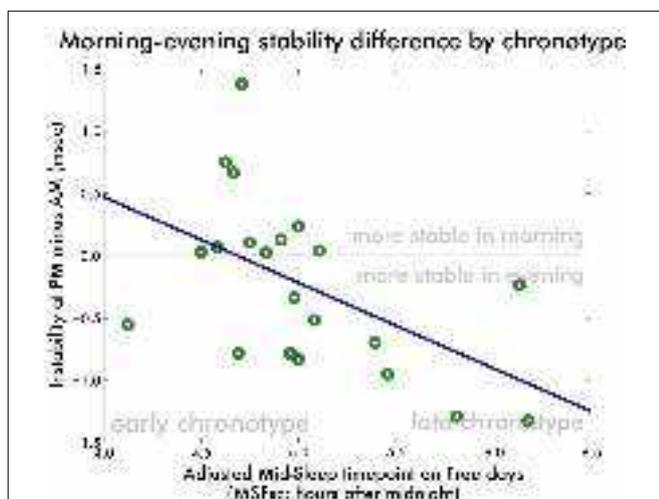
median of these for each participant and condition. An ANOVA with this keystroke velocity as dependent variable revealed no main effect of hand or chronotype [both  $F_{(1, 19)} < 3.00$ ,  $p > 0.14$ ]. However, there was a main effect of scale direction [ $F_{(1, 19)} = 5.69$ ,  $p < 0.03$ ,  $\eta_G^2 = 0.01$ ], revealing that outward scales were played louder ( $M = 85.4$ ,  $SD = 4.1$  MIDI units) than inward scales ( $M = 84.6$ ,  $SD = 4.4$  MIDI units). Furthermore, there was a main effect of recording [ $F_{(1, 19)} = 10.21$ ,  $p = 0.005$ ,  $\eta_G^2 = 0.03$ ], which indicated that keystroke velocity during AM recordings was slower ( $M = 84.2$ ,  $SD = 4.4$  MIDI units) than using the PM recordings ( $M = 85.8$ ,  $SD = 4.3$  MIDI units). There was a trend for an interaction between pianist chronotype and recording [ $F_{(1, 19)} = 4.22$ ,  $p = 0.05$ ,  $\eta_G^2 = 0.01$ ]. This revealed that the keystroke velocity increase between the AM and PM recordings was negligible for early chronotype pianists ( $M = 0.5$ ,  $SD = 2.5$  MIDI units) but considerable for late chronotype pianists ( $M = 2.5$ ,  $SD = 2.1$  MIDI units).

We continued to analyse the keystroke velocity unevenness (the median of the SD of the keystrokes in each scale run). We found no main effect of chronotype [ $F_{(1, 19)} = 1.26$ ,  $p = 0.27$ ] but there was a main effect of hand [ $F_{(1, 19)} = 5.14$ ,  $p = 0.04$ ,  $\eta_G^2 = 0.02$ ], revealing that keystroke velocities were more even in the right ( $M = 5.8$ ,  $SD = 1.5$  MIDI units) than in the left hand ( $M = 6.2$ ,  $SD = 1.3$  MIDI units). There were no other main effects [both  $F_{(1, 19)} < 0.87$ ]. There was an interaction effect of hand and direction [ $F_{(1, 19)} = 24.69$ ,  $p < 0.001$ ,  $\eta_G^2 = 0.07$ ], revealing that for the left hand the keystroke velocities of the outward scales were more even than the inward scales, whereas for the right hand the opposite was true. Crucially, there was an interaction between pianist chronotype and recording time-point [ $F_{(1, 19)} = 5.67$ ,  $p = 0.03$ ,  $\eta_G^2 = 0.02$ ], revealing that keystroke velocities of pianists of the early chronotype were more even in the AM recordings than in the PM recordings, and vice versa for the late chronotype pianists. There were no other interactions [all  $F_{(1, 19)} < 1.83$ ,  $p > 0.19$ ]. We emphasize that these results in keystroke velocity unevenness should be interpreted with caution, because of the differences in baseline velocity between the groups, as well as the interaction with chronotype mentioned above.

### DISCUSSION

We investigated the influence of chronotype on fine motor performance, taking playing of musical scales in pianists as an example. Piano students were recorded playing scales in the morning and evening. The participant pool was divided into early and late sleepers in a way that no differences in musical training, age, or gender between the groups occurred. These two groups were shown to be comparable in overall temporal precision in scale playing and also revealed no differences the temporal unevenness of scale playing between the morning and evening recording sessions.

However, the unevenness metric assesses overall variability of scale timing without taking into account that part of the variability is consistent between trials. Indeed, it has been shown that some of the notes are consistently late or early (van Vugt et al., 2012) in such a way that a highly individual temporal deviation pattern appears. This trace is mostly due to differential biomechanical properties of the motor system of individual pianists



**FIGURE 3 | Difference in playing instability between the morning and evening recordings as a function of pianist chronotype.** We have collapsed the two hands and two directions to yield a single data point for each participant. Positive instability differences denote more stable playing in the morning. A negative correlation is observed, indicating that the later the chronotype, the more stable the scale playing is in the evening relative to the morning.

(van Vugt et al., 2013). Therefore, we expected that this trace would not be susceptible to circadian fluctuations, since basic properties of the motor system such as muscle- and joint materials remain the same throughout the day. This prediction was largely supported, as at most a slight tendency was found for circadian variation in irregularity. The residual variability, that is, the variability that diverged from this individual trace (instability), revealed a different picture. Although there was no overall difference in timing stability between the pianists, an interaction was found between the recording time and the pianists' chronotype. This revealed that late sleepers' timing was more stable in the evening than in the morning, whereas early sleepers exhibited no difference between the two recordings. Our result can therefore be added to a growing list of aspects of human performance that are subject to circadian fluctuations (Kleitman et al., 1938; Atkinson et al., 1993; Coldwells et al., 1994; Gauthier et al., 1996; Monk and Kupfer, 2000; Dosseville et al., 2002; Van Eekelen and Kerkhof, 2003). In our context, it is interesting to notice that even long-term, massed practice of these expert musicians has not been able to wash out circadian fluctuations in performance. As such, we suggest that such fluctuations are much more deeply embedded in the human motor system than might otherwise be assumed.

The question that remains open, is why the early chronotype group did not show a difference in instability between the recordings (as the late chronotype did). On the other hand, the regression analysis on both groups combined showed a significant correlation between chronotype and morning-evening instability difference. An explanation might be that our participant pool did not include sufficiently early chronotypes. Indeed, the early chronotype mid-sleep time point (MSFsc) was not early compared to a reference population reported by Roenneberg et al. (2004). Therefore, a future study could include earlier chronotypes and might reveal the recording effect not found in this study.

A secondary thread of analyses of keystroke velocities revealed that in the evening, keystroke velocities were higher than in the morning, and there was a trend for this inequality to be different between the two chronotypes. This result may be interpreted in analogy with previous findings on physical force that underlie a circadian fluctuation with a minimum in the morning and a maximum in the afternoon or early evening. Left hand maximal grip strength (Atkinson et al., 1993) was, for example, significantly higher in the evening than in the morning. A similar finding was reported for left elbow flexor torque (Gauthier et al., 1996). Although these results were reported for completely different settings it is possible that fluctuations of physical force influence piano playing: pianists may tend to exert higher forces while playing the piano as soon as the maximum available forces are higher—with the result of higher key velocities in the evening. Finally, keystroke velocity analyses revealed that early chronotypes were more even in the morning and late chronotypes more even in the evening. This finding, although in line with our hypothesis of a chronotype-specific optimal playing time in the daily cycle, should be interpreted with caution. The reason is that the baseline velocities were different between the two measurements and differentially so for the two chronotypes.

It is important to realize that our result is based on the mid-sleep time-point on free days adjusted for individual average sleep need accumulated on work days (MSFsc). The measurements themselves, however, took place on work days. In this way, we have eliminated possible influences of larger variations in sleeping behavior on free days (Roenneberg et al., 2003b). However, since many participants reported a number of 7 work days per week, the aforementioned adjustment was relevant in a minority of participants. As a consequence, the influence of the recordings being carried out on work days may be limited in our study sample. A limitation of the present study is that we did not collect data on the participants' sleep quantity and quality of night before.

On the level of the scale playing analysis methodology, we found that the left hand played the scales less evenly, and in particular with greater instability. The amount of discarded scale material (due to errors) was not different between the two hands. Our result replicates the previous finding that left hand playing was shown to be less evenly in a scale playing task, both for right- and left-handed pianists (Kopiez et al., 2011).

How salient are the differences we present here? We feel that the present result, although statistically reliable, is subtle. Timing differences were in the order of milliseconds. Previous research showed that even expert musicians are insensitive to timing irregularities below approximately 10 ms unevenness (van Vugt et al., 2013), suggesting that concert audience's appreciation is not likely to be much influenced by the performer's chronotype. However, the possibility remains that when the pianists' capacities are taken to the limit, such as in playing a challenging piano concerto over sustained (multi-hour) periods of time, these circadian fluctuations become perceptible and a determining factor in the appreciation of the performance.

A future study might aim at replicating our result but with a larger range of chronotypes. Our study has not specifically recruited extreme chronotypes but instead opted for a representative sample of the music student population. The mid-sleep time point (MSFsc) values of the students in our sample were approximately within  $\pm 1$  SD of the range of the mean MSFsc values yielded in a large population for this age group (Roenneberg et al., 2004) (see **Figure A1**). Our prediction is that more extreme chronotypes will show a greater difference in performance between morning and evening recordings. Additional questions for future research are whether the size of the instability difference between morning and evening recordings is different for the early and late sleepers and whether the deficit in performance skills due to circadian fluctuation may be reduced through a shift of the sleep phase prior to a performance at a chronobiologically unfavorable time of the day.

## ACKNOWLEDGMENTS

This work was supported by the EBRAMUS, European Brain and Music Ph.D. Grant to Floris T. Van Vugt (ITN MC FP7, GA 238157). The authors also wish to thank Prof. Dr. Till Roenneberg and Dr. Marc Bangert for discussions during the earlier stages of this research. Also, we are indebted to Prof. Dr. Till Roenneberg for kindly providing us with the Munich Chrono-Type Questionnaire.

## REFERENCES

- Aschoff, J. (1955). Der tagesgang der Körpertemperatur beim Menschen. *J. Mol. Med.* 33, 545–551.
- Atkinson, G., Coldwells, A., Reilly, T., and Waterhouse, J. (1993). A comparison of circadian rhythms in work performance between physically active and inactive subjects. *Ergonomics* 36, 273–281. doi: 10.1080/00140139308967882
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384. doi: 10.3758/BF03192707
- Benoit, O., Foret, J., Merle, B., and Bouard, G. (1981). Diurnal rhythm of axillary temperature in long and short sleepers: effects of sleep deprivation and sleep displacement. *Sleep* 4, 359–365.
- Coldwells, A., Atkinson, G., and Reilly, T. (1994). Sources of variation in back and leg dynamometry. *Ergonomics* 37, 79–86. doi: 10.1080/00140139408963625
- Dosseville, F., Moussay, S., Larue, J., Gauthier, A., and Davenne, D. (2002). Physical exercise and time of day: influences on spontaneous motor tempo. *Percept. Mot. Skills* 95, 965–972.
- Gauthier, A., Davenne, D., Martin, A., Cometti, G., and Hoecke, J. V. (1996). Diurnal rhythm of the muscular performance of elbow flexors during isometric contractions. *Chronobiol. Int.* 13, 135–146. doi: 10.3109/07420529609037077
- Hasher, L., Zacks, R. T., and Rahhal, T. A. (1999). Timing, instructions, and inhibitory control: some missing factors in the age and memory debate. *Gerontology* 45, 355–357. doi: 10.1159/000022121
- Horne, J. A., and Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int. J. Chronobiol.* 4, 97–110.
- Intons-Peterson, M. J., Rocchi, P., West, T., McLellan, K., and Hackney, A. (1999). Age, testing at preferred or nonpreferred times (testing optimality), and false memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 23. doi: 10.1037/0278-7393.25.1.23
- Jabusch, H.-C., Vauth, H., and Altenmüller, E. (2004). Quantification of focal dystonia in pianists using scale analysis. *Mov. Disord.* 19, 171–180. doi: 10.1002/mds.10671
- Jasper, I., Roenneberg, T., Häußler, A., Zierdt, A., Marquardt, C., and Hermsdörfer, J. (2010). Circadian rhythm in force tracking and in dual task costs. *Chronobiol. Int.* 27, 653–673. doi: 10.3109/07420521003663793
- Kantermann, T., Forstner, S., Halle, M., Schlangen, L., Roenneberg, T., and Schmidt-Trucksäss, A. (2012). The stimulating effect of bright light on physical performance depends on internal time. *PLoS ONE* 7:e40655. doi: 10.1371/journal.pone.0040655
- Kleitman, N., Titelbaum, S., and Feiveson, P. (1938). The effect of body temperature on reaction time. *Am. J. Physiol. Legacy Content* 121, 495–501.
- Kopiez, R., Jabusch, H.-C., Galley, N., Homann, J.-C., Lehmann, A. C., and Altenmüller, E. (2011). No disadvantage for left-handed musicians: the relationship between handedness, perceived constraints and performance-related skills in string players and pianists. *Psychol. Music* 40, 357–384. doi: 10.1177/0305735610394708
- Matchock, R. L., and Mordkoff, J. T. (2009). Chronotype and time-of-day influences on the alerting, orienting, and executive components of attention. *Exp. Brain Res.* 192, 189–198. doi: 10.1007/s00221-008-1567-6
- May, C. P., Hasher, L., and Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychol. Sci.* 4, 326–330. doi: 10.1111/j.1467-9280.1993.tb00573.x
- Morrow, M., Spoelstra, K., and Roenneberg, T. (2005). The circadian cycle: daily rhythms from behaviour to genes. *EMBO Rep.* 6, 930–935. doi: 10.1038/sj.embor.7400541
- Monk, T. H., and Kupfer, D. J. (2000). Circadian rhythms in healthy aging-effects downstream from the pacemaker. *Chronobiol. Int.* 17, 355–368. doi: 10.1081/CBI-100101051
- Münste, T. F., Altenmüller, E., and Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* 3, 473–478.
- Peres, I., Vetter, C., Blautzik, J., Reiser, M., Pöppel, E., Meindl, T., et al. (2011). Chronotype predicts activity patterns in the neural underpinnings of the motor system during the day. *Chronobiol. Int.* 28, 883–889. doi: 10.3109/07420528.2011.619084
- Petros, T. V., Beckwith, B. E., and Anderson, M. (1990). Individual differences in the effects of time of day and passage difficulty on prose memory in adults. *Br. J. Psychol.* 81, 63–72. doi: 10.1111/j.2044-8295.1990.tb02346.x
- Reilly, T., Atkinson, G., Edwards, B., Waterhouse, J., Farrelly, K., and Fairhurst, E. (2007). Diurnal variation in temperature, mental and physical performance, and tasks specifically related to football (soccer). *Chronobiol. Int.* 24, 507–519. doi: 10.1080/07420520701420709
- Roenneberg, T., Daan, S., and Mellow, M. (2003a). The art of entrainment. *J. Biol. Rhythms* 18, 183–194. doi: 10.1177/0748730403018003001
- Roenneberg, T., Wirz-Justice, A., and Mellow, M. (2003b). Life between clocks: daily temporal patterns of human chronotypes. *J. Biol. Rhythms* 18, 80–90. doi: 10.1177/0748730402239679
- Roenneberg, T., Kuehnle, T., Pramstaller, P. P., Ricken, J., Havel, M., Guth, A., et al. (2004). A marker for the end of adolescence. *Curr. Biol.* 14, R1038–R1039. doi: 10.1016/j.cub.2004.11.039
- Roenneberg, T., and Mellow, M. (2003). The network of time: understanding the molecular circadian system. *Curr. Biol.* 13, R198–R207. doi: 10.1016/S0960-9822(03)00124-6
- Tamm, A. S., Lagerquist, O., Ley, A. L., and Collins, D. F. (2009). Chronotype influences diurnal variations in the excitability of the human motor cortex and the ability to generate torque during a maximum voluntary contraction. *J. Biol. Rhythms* 24, 211–224. doi: 10.1177/0748730409334135
- Van Eekelen, A. P. J., and Kerkhof, G. A. (2003). No interference of task complexity with circadian rhythmicity in a constant routine protocol. *Ergonomics* 46, 1578–1593. doi: 10.1080/0014013031000121598
- van Vugt, F. T., Jabusch, H.-C., and Altenmüller, E. (2012). Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Front. Psychol.* 3:495. doi: 10.3389/fpsyg.2012.00495
- van Vugt, F. T., Jabusch, H.-C., and Altenmüller, E. (2013). Individuality that is unheard of: systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint. *Front. Psychol.* 4:134. doi: 10.3389/fpsyg.2013.00134
- Wagner, C. (1971). “The influence of the tempo of playing on the rhythmic structure studied at pianist's playing scales,” in *Medicine and Sport Biomechanics II*, eds J. Vredenburg and J. Wartenweiler (Basel: Karger), 129–132.
- West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., and Stuss, D. T. (2002). Effects of time of day on age differences in working memory. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 57, P3–P10. doi: 10.1093/geronb/57.1.P3
- Wisser, H., and Breuer, H. (1981). Circadian changes of clinical chemical and endocrinological parameters. *J. Clin. Chem. Clin. Biochem.* 19, 323–337.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

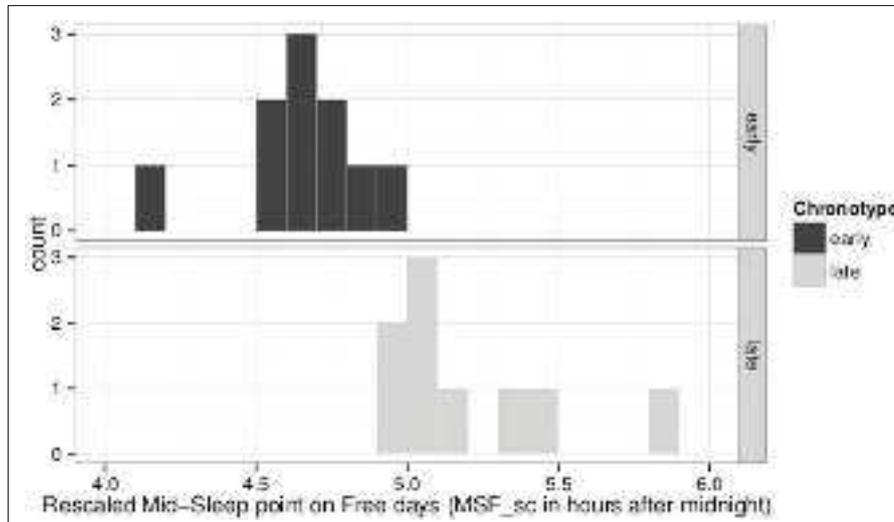
Received: 28 February 2013; paper pending published: 14 April 2013; accepted: 17 June 2013; published online: 09 July 2013.

Citation: Van Vugt FT, Treutler K, Altenmüller E and Jabusch H-C (2013) The influence of chronotype on making music: circadian fluctuations in pianists' fine motor skills. *Front. Hum. Neurosci.* 7:347. doi: 10.3389/fnhum.2013.00347

Copyright © 2013 Van Vugt, Treutler, Altenmüller and Jabusch. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

**APPENDIX**

**DISTRIBUTION OF MID-SLEEP TYPE POINTS**



**FIGURE A1 |** Distribution of the adjusted mid-sleep time point on free days (MSFsc) for the early and late chronotype groups.

**QUESTIONNAIRE**

Below are the excerpts from the questionnaire that we used to establish participant’s mid-practice time and accumulated practice time. The questions are translated from German.

- Surname: \_\_\_\_\_ First name: \_\_\_\_\_
- Date of birth: \_\_\_\_\_ Gender (m/f): \_\_\_\_\_
- Phone: \_\_\_\_\_ E-mail: \_\_\_\_\_
- Study path: \_\_\_\_\_ Total number of semesters: \_\_\_\_\_
- Today’s date: \_\_\_\_\_

(1) How old were you when you started to play the piano?  
 \_\_\_\_\_ years.

(2) Please indicate the amount of your average daily practice time in hours for the respective age segment.

Age (years)	0–10	11–15	16–20	21–25	26–30	31–35
-------------	------	-------	-------	-------	-------	-------

Hours

(3) Please indicate the amount of your current average practice time in hours for the respective times of the day.

Time of the day	0–3 a.m.	3–6 a.m.	6–9 a.m.	9–12 a.m.	12–15 p.m.	15–18 p.m.	18–21 p.m.	21–24 p.m.
-----------------	----------	----------	----------	-----------	------------	------------	------------	------------

Hours (maximum: 3 h per cell)

## **VIII.2. Musician's dystonia in pianists: long-term evaluation of retraining and other therapies.**

---

This study reveals how musician's dystonia is a problem of timing in production of over-learned movement sequences (such as musical scales). We show that changes in sensory feedback as implemented in behavioural retraining therapy can provide an effective means of therapy.













## **IX. Manuscript listing**

**Study II.2** (p. 92)

van Vugt FT, & Tillmann B ([in preparation](#)). Thresholds of auditory-motor coupling measured with a simple task in musicians and non-musicians: Was the sound simultaneous to the key press?

**Study II.3** (p. 116)

van Vugt FT, & Tillmann B ([in preparation](#)). Auditory feedback benefits short term motor sequence learning.

**Study III.2** (p. 178)

van Vugt FT, Kafczyk T, Kuhn W, Rollnik J, Tillmann B, & Altenmüller ([in preparation](#)). Random delay boosts musical fine motor recovery after stroke.

**Study III.3** (p. 220)

van Vugt FT, Ritter J, Rollnik J, & Altenmüller ([in preparation](#)). Music-supported motor training after stroke reveals no superiority of synchronisation in group therapy.

**Study IV.2** (p. 260)

van Vugt FT, Jabusch HC, & Altenmüller E (2012). Fingers phrase music differently: trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Frontiers in Psychology*. 3:495. doi: 10.3389/fpsyg.2012.00495.

**Study IV.3** (p. 270)

van Vugt FT, Jabusch HC, & Altenmüller E (2013). Individuality that is unheard of: systematic temporal deviations in scale playing leave an inaudible pianistic fingerprint. *Frontiers in Psychology*. 4:134. doi: 10.3389/fpsyg.2013.00134

**Study IV.4** (p. 282)

van Vugt FT, Furuya S, Vauth H, Jabusch HC, & Altenmüller E (in revision). Spatial and temporal symmetries of motor primitives in skilled piano performance at different tempi.

**Study VII.1** (p. 372)

van Vugt FT, Schuepbach M, Altenmüller E, Bardinet E, Yelnik J, & Haelbig, TD (2013). Effects of dopaminergic and subthalamic stimulation on musical performance. *Journal of Neural Transmission*. 120(5), 755–759. doi:10.1007/s00702-012-0923-7

**Study VIII.1** (p. 380)

van Vugt FT, Treutler K, Altenmüller E, & Jabusch HC (2013). The influence of chronotype on making music: Circadian fluctuations in pianists' fine motor skills. *Frontiers in Human Neuroscience*. 7:347. doi: 10.3389/fnhum.2013.00347

**Study VIII.2** (p. 390)

van Vugt FT, Boullet L, Jabusch HC, & Altenmüller E (accepted). Musician's dystonia in pianists: long-term evaluation of retraining and other therapies. *Parkinsonism and Related Disorders*. doi: 10.1016/j.parkreldis.2013.08.009.

**International conference presentations associated with the thesis:**

van Vugt FT and Tillmann B (2013). Thresholds of auditory-motor coupling measured with a simple task in musicians and non-musicians: Was the sound simultaneous to the key press? [Progress in Motor Control: IX](#), 14-16 July 2013. Montreal, Canada.

van der Steen MC, van Vugt FT, Keller PE, Altenmueller E (2013). Separating perception and production abilities in auditory-motor processing of musician's dystonia patients. [Progress in Motor Control: IX](#), 14-16 July 2013. Montreal, Canada.

Jabusch HC, van Vugt FT, Bangert M, Treutler K, and Altenmuller E (2013). Piano playing and chronotype: Chronobiological influences on sensorimotor precision in pianists. [International Symposium of Performance Science](#). Vienna, 28-31 August 2013.

van Vugt FT, Kafczyk T, Kuhn W, Rollnik JD, Tillmann B, Altenmueller E. (2013) Random delay boosts musical fine motor recovery after stroke. [11th meeting of the French Neuroscience Society](#), Lyon, 21-24 May 2013.

van der Steen MC, van Vugt FT, Keller PE, Altenmueller E (2013). Perception and production abilities of musician's dystonia patients in auditory-motor processing. [Second International Congress on Treatment of Dystonia](#). Hannover, 8-11 May 2013.

van Vugt FT, Kuhn W, Rollnik J, & Altenmueller E. (2012). Random delay boosts musical fine motor recovery after stroke. [12th International Conference on Music Perception and Cognition: ICMPC-ESCOM 2012](#), 23-28 July 2012, Tessaaloniki, Greece.

Haelbig T, Altenmueller E, Schuepbach M, van Vugt FT (2012). Effects of dopaminergic and subthalamic stimulation on musical performance parameters: Dissociating timing, intonation, articulation and emotionality. [16th International Congress on Parkinson's](#)

Disease and Movement Disorders. 17-21 June 2012, Dublin, Ireland.

van Vugt FT, Haelbig T, Schuepbach M, Buttkus F, & Altenmueller E. (2011). The musician doctor: A musical evaluation of treatments for movement disorders. [International Symposium on Performance Science](#) 2011, 24-27 August, Toronto, Canada.

van Vugt FT, Cheng F, Jabusch HC, & Altenmueller E. (2011). Fine temporal deviations in scale playing reveal motor program chunking. [The Neurosciences and Music](#), Edinburgh, Scotland, UK, 9 - 12 June 2011.

Sounds on time - Manuscript listing

## **X. Acknowledgements**

### Words of thanks

For in truth it is life that gives unto life - while you,  
who deem yourself a giver, are but a witness.

– Kahlil Gibran, *The Prophet*

A thesis is a big piece of work but it is not the work of me alone. It is humbling and inspires great gratitude to realise how none of this could have happened if it were not for a great number of extraordinary people along the way. My thanks could never do justice to each and every one, but I wanted to specifically extend my gratitude to the following people.

Prof. Eckart Altenmüller, your relentless enthusiasm taught me never to forget that we have to enjoy doing science. You gave me the freedom to work on the questions that I wanted to work on, and always gave your enthusiasm and respect, for which I am deeply grateful. I appreciate our friendship and look forward to playing music together again. I thank Dr. Barbara Tillmann who has shown me the importance of being thorough, attentive and precise. What I've learned from you was fundamental for each and every study in this thesis. I owe a particular word of thanks to Prof. Patrick Cavanagh, who has taught me everything in science from the ground up. I can only hope to live up to all of what you taught me one day. Prof. Hans-Christian Jabusch, with an aviator's eye for detail, I thank you for always making scientific collaboration enjoyable. Dr. Shinichi Furuya, by believing in me you helped me to believe in myself; thank you for becoming a good and valued friend. Arigatoo! Philippe Albouy, I thank you for always being unreasonably admiring (which is reciprocal), and for many enjoyable hours of squash. A word of thanks to Dr. Michael Grossbach, who was always available for scientific discussion. Marta Beauchamp,

Maité Castrò, Alexandra Corneyllie, Daniel Scholz, Yohana Lévêque, Tatiana Selchenkova, Marieke van der Steen, and Britta Westner: thanks for being fun office mates in Hannover and Lyon and great collaborators. Yousreya Pölkner, thanks for making the IMMM such a friendly (and well-organised!) place. I thank the excellent clinicians who did almost all of the patient work reported in this thesis much better than I could have: Wolfgang Kuhn, Tom Kafczyk and Juliane Ritter.

Meaghan Fowle, it meant a lot to me to share the joys and challenges of this thesis with you. I sometimes feel the sacrifices you made for this thesis exceed my own. Meaghan was also an invaluable language editor. Vera Scepanovic, thank you for offering me a beer on the beach (and for beautiful artwork on page 281 with our friend Mariana Fulgueiras). My UCLA linguistics cohort deserves a special thanks for being inspiring and engaging. Thank you to all members of the Lyon Neuroscience Research Center as well as all members of the Hannover Institute of Music Physiology for many enjoyable hours together. Many thanks to the supervisors of the EBRAMUS network for support, and to the students for sharing our Ph.D. ups and downs. EBRAMUS kids are all right!

No word of thanks I could write here could possibly do justice to what I owe to my parents, my sisters and all my family for always supporting me. Thank you.

Sogyal Rinpoche I thank and all other masters, for inspiring me and all beings, and thanks to all the friends on the path.

In particular, I want to thank all those who have believed in me, especially in those moments when I didn't myself; you have given me the greatest gift of all.

Nella vita degli imperatori c'è un momento, che segue [...] alla malinconia e al sollievo di sapere che presto rinunceremo a conoscerli e a comprenderli; un senso come di vuoto [...]: è il momento disperato in cui si scopre che quest'impero che ci era sembrato la somma di tutte le meraviglie è uno sfacelo senza fine né forma [...] che il trionfo sui sovrani avversari ci ha fatto eredi della loro lunga rovina.

- Italo Calvino, *Città Invisibili*



