



HAL
open science

Four essays on the psychological determinants of risk-taking, education and economic performance

Noémi Berlin

► **To cite this version:**

Noémi Berlin. Four essays on the psychological determinants of risk-taking, education and economic performance. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2013. English. NNT : 2013PA010053 . tel-00984291

HAL Id: tel-00984291

<https://theses.hal.science/tel-00984291v1>

Submitted on 28 Apr 2014

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.

UNIVERSITÉ PARIS 1 - PANTHÉON SORBONNE
U.F.R DE SCIENCES ECONOMIQUES

Numéro attribué par la bibliothèque

|||||

Thèse pour le doctorat de Sciences Economiques

soutenue publiquement par

Noémi BERLIN

le Mardi 3 Décembre 2013

**FOUR ESSAYS ON THE PSYCHOLOGICAL DETERMINANTS OF
RISK-TAKING, EDUCATION AND ECONOMIC PERFORMANCE**

Directeur de thèse

Louis LÉVY-GARBOUA Professeur à l'Université Paris 1 Panthéon-Sorbonne

Jury

Rapporteurs : Susan LAURY Professeure à Georgia State University
Luis SANTOS PINTO Professeur à HEC Lausanne
Président : Jean-Christophe VERGNAUD Directeur de Recherche CNRS - CES
Examineurs : Todd LUBART Professeur à l'Université Paris-Descartes
Marie-Claire VILLEVAL Directrice de Recherche CNRS - GATE

L'Université Paris I Panthéon-Sorbonne n'entend donner aucune approbation, ni improbation aux opinions émises dans cette thèse; elles doivent être considérées comme propres à leur auteur.

Remerciements

Les remerciements, presque un chapitre à part entière, sont toujours lus attentivement. Lorsqu'on arrive au moment de les rédiger, cela signifie bel et bien la fin. Ils me permettent alors de citer toutes les personnes qui m'ont le plus soutenue pendant ces quatre années longues, tumultueuses mais aussi tellement stimulantes et passionnantes.

Tout d'abord, je remercie mon directeur, Louis Lévy-Garboua. Il a cru en moi et m'a donné la chance de construire ces quatre années de travail. Plus qu'un directeur de thèse, il est un entraîneur! De nature assez stressée, je me suis souvent retrouvée en face de lui désespérée, perdue, paniquée et pleine de doutes. Sa qualité fut de toujours trouver le mot juste pour me remettre d'aplomb et de toujours trouver une solution. J'admire son optimisme qui m'a souvent été d'une grande aide, sa clairvoyance sur l'ensemble de mon travail. Ma liberté a été telle que j'ai pu entreprendre différents projets, et le soutien, même de loin, n'a jamais manqué.

Merci aux membres du jury: Susan Laury, Todd Lubart, Luis Santos Pinto, Jean-Christophe Vergnaud et Marie-Claire Villeval, qui ont accepté d'en faire partie et qui se sont déplacés, parfois de loin. Leurs remarques et critiques ont été d'une grande richesse et m'ont permis d'améliorer grandement ce manuscrit. Je ferai ensuite de mon mieux pour intégrer le reste de leurs commentaires afin de perfectionner les articles de cette thèse mais j'en prendrai aussi note pour mes travaux futurs.

La thèse au quotidien c'est évidemment la Maison des Sciences Economiques, lieu que l'on déteste à la fin du Master et qu'on finit par apprécier malgré soi pendant son doctorat. Le couloir du quatrième aka UG4 est rempli de conseils, d'aides et de remarques indispensables. Je remercie donc Margherita Comola, Stéphane Gauthier, Nicolas Jacquemet, David Margolis, Benoît Rapoport, Antoine Terracol et Jean-Christophe Vergnaud. Merci encore à Antoine, Stéphane et Jean-Philippe Tropéano de m'avoir confié leurs TDs. Merci aussi à Claude Montmarquette et Nathalie Viennot.

Je remercie tout le personnel administratif et informatique, avec une pensée particulière vers Laurie Altemaire, Elda André (sans qui de nombreux doctorants seraient perdus), Loïc Sorel, Tonia Lastapis et Stéphane Brice. Merci à Joël et tout le personnel de l'accueil.

Mes compétences en programmation étant limitées, heureusement que Maxim Frolov

était là pour programmer la plupart des expériences de cette thèse.

Merci aux thésards qui m'ont accueillie: Marie-Pierre, avec qui j'ai eu l'occasion d'écrire un des mes articles et qui n'a jamais manqué de m'aider. Clémence et Gaëlle grâce à qui je suis devenue une doctorante informée et introduite aux méandres du monde de la recherche. J'ai gagné d'innombrables heures en leur posant mes montagnes de questions sur nos chers et tendres logiciels Latex et Stata. Ce qui m'en vient à remercier Adam pour toutes les commandes Stata qu'il a pu m'apprendre ainsi que pour sa bonne humeur et ses blagues quotidiennes! Merci à Nicolas, Sébastien et Batool avec qui j'ai partagé le bureau 402 pendant quatre années. Je n'oublierai jamais la danse du Sabar si bien mise en oeuvre par Omar. Camille, Djamel et Emma, merci de m'avoir permis l'accès au monde obscur des macroéconomistes. Merci à tous les autres doctorants de l'UG4: Alexandra, Chaimaa, Claire, Dorian, Juliette, Théa, Thomas et Rawa.

Je remercie aussi le collège Saint-Exupéry de Rosny-sous-Bois ainsi que José de Valverde. Trouver un collègue qui accepte de nous ouvrir ses portes est un périple rempli d'embûches. Je n'aurais pas pu mener à bien cette thèse sans leur accord.

Je n'oublie pas la team de l'institut de Psychologie de Paris-Descartes. Grâce à Jean-Louis j'ai appris ce qu'était un alpha, primordial! Sa réactivité aux mails m'a toujours épatée et ce fut un réel plaisir de travailler avec lui. Merci à Franck, Maud, Todd et Xavier pour leurs aides variées. Je suis ravie d'avoir pu mettre un pied dans le monde mystérieux de la psychologie. J'ai pu, grâce à eux, participer à l'entente si tumultueuse mais exaltante entre les économistes et les psychologues.

Un remerciement particulier pour Rémi, doctorant flambant neuf que j'ai pu mettre à contribution durant ma dernière ligne droite, ainsi que pour Marco pour toute l'aide (et ses gâteaux) qu'il m'a apportée.

Lorsque je n'étais pas au labo, il y a ces personnes irremplaçables qui ne voyaient pas ce quotidien mais qui m'ont soutenue, encouragée, écoutée et m'ont permis d'oublier ma thèse le temps de soirées, vacances et cafés.

Je remercie les vieilles du temple: Fred, Lydia, Marie, Maxime, Nicolas, Thomas R.. Merci à Lydia d'avoir corrigé mon anglais (qui en avait bien besoin) à plusieurs reprises. Je remercie Marie pour sa folie, sa bonne humeur et sa compréhension.

Un immense remerciement envers Carla, Hélène, Lorraine et Lucie qui après tout

ce temps sont toujours dans ma vie.

Nos chemins se sont séparés au tout début de cette thèse, mais je remercie Léo pour ses présences éparses mais non négligeables.

Jérôme, Thomas P.: un super gigantesque MERCI. Chacun, à votre façon, vous avez toujours été là!

J'ai longuement réfléchi pour savoir où les introduire, mais je crois que leur place dépasse largement les frontières de la MSE: Anna et Léontine. La dernière année de thèse étant la plus dure, je ne sais pas ce que j'aurais fait sans elles. Elles étaient...mon brumisateuse pendant une canicule! Cette brise qui rend les journées supportables, sans qui le quotidien au labo n'aurait pas été le même. Merci merci merci! Elles sont mes plus belles trouvailles de cette aventure.

Nous avons fait une bonne partie de nos études ensemble, nous avons commencé notre thèse ensemble, nous la finissons ensemble, à quelques jours près. Nos humeurs en dents de scie étaient parfaitement complémentaires et nous ont permis de garder le cap! C'était une chance d'avoir quelqu'un qui vivait les mêmes joies et déceptions à la minute près. Lisa, 1000 mercis pour toutes les fois où on s'est retrouvé en haut ou en bas de notre fameuse vague!

Une dédicace spéciale à la clique: Antonina, Daphné, Dimitri, Joyce, Lizhen, Marion, Thibault et Xiaou. Plus de 10 ans que l'on se suit, ils ont toujours été là. M.E.R.C.I. J'ai toujours essayé d'écouter attentivement les conseils si pertinents d'Antonina. Merci à Daphné d'avoir toujours été présente et d'avoir participé à mes expériences! Je remercie Dimitri pour son humour décompressant. Joyce est la seule qui comprenait vraiment ce que je faisais. Elle m'a suivie de près pendant tout ce temps, m'a secouée à plusieurs reprises, a relu et corrigé d'innombrables pages, un soutien primordial. Lizhen, merci d'avoir toujours essayé de comprendre ce que je faisais et pour tes nombreux messages encourageants. J'espère que Thibault et Judith ne regretteront pas l'achat de la Revue Eco (quelle surprenante attention!). La force de Xiaou m'a toujours permis de garder les pieds sur terre. La clique forever, vous êtes mes âmes sœurs!

L'idée de faire une thèse ne m'est pas venue naturellement, n'a jamais été une vocation et je ne remerciais jamais assez mon père de m'avoir poussée à m'engager dans cette aventure vertigineuse, mais tellement excitante. Son soutien était omniprésent

et sa fierté inconditionnelle, ce qui représente bien un moteur de tous les jours. Je remercie évidemment ma mère qui n'a jamais cessé de m'encourager, de m'épauler et m'a surtout aidée à relativiser. Comme mon père, son objectivité envers moi n'existe tout simplement pas. Je ne peux que bien faire. Alors, lorsque je perdais pieds, c'était toujours, sans l'assumer, que j'attendais ses paroles énergisantes.

Au tour de mes grands-parents, Nathan et Marie. Cette thèse m'a rendue plus absente que d'habitude, mais aucune des mes visites et appels n'ont manqués de cette question "Alors le boulot, ça avance?". J'admire mon grand-père qui a toujours essayé de se plonger dans mes papiers malgré une grande frustration, et je le vois brandir fièrement ma première publication dans la Revue Economique. Sans ma grand-mère, ces quatre années, comme avant cette thèse, auraient été bien pauvres en renfort et réconfort. Elle est tout simplement inestimable.

Merci à mes frères et à ma sœur. Sans eux je ne serais bizarrement pas grand chose. A la question incessante de mes frères "et alors tu fais quoi après? tu pars en post-doc?" jusqu'aux histoires contées par ma sœur constamment enjouée, les voir et leur parler ont toujours amélioré mes journées.

À tous ceux que je n'ai pas mentionnés mais qui le méritent tout autant: Merci!

On dit qu'une thèse est un travail purement individuel rempli de bons et de mauvais moments. Il faut apprendre seul à se relever et à se motiver. Mais on ne réussit pas à la finir si on n'est pas entouré. J'ai eu la chance d'être très soutenue et aidée par toutes ces personnes citées. Sans eux c'est certain, ce manuscrit n'aurait sûrement jamais existé.

À mes grand-parents, mes parents, mes frères et ma soeur.

Avertissement

Mis à part l'introduction et la conclusion de cette thèse, les différents chapitres sont issus d'articles de recherche rédigés en anglais et dont la structure est autonome. Par conséquent, des termes "papier" ou "article" y font référence, et certaines informations, notamment la littérature, sont répétées d'un chapitre à l'autre.

Notice

Except the general introduction and conclusion, all chapters of this thesis are self-containing research articles. Consequently, terms "paper" or "article" are frequently used. Moreover, some explanations, like corresponding literature, are repeated in different places of the thesis.

Table of Contents

Résumé	1
Abstract	3
General Introduction	5
0.1 The empirical method	9
0.2 Risk preferences	10
0.3 Confidence	14
0.4 Intrinsic Variables	17
0.4.1 Personality traits	17
0.4.2 Gender	19
0.5 Creativity Potential	21
0.6 Outline of the dissertation	23
I Confidence and Economic Behaviors	29
1 Confidence, Aspiration and Performance	31
1 Introduction	32
2 The experiment	37
2.1 The design	37
2.2 Confidence elicitation	40
2.3 Descriptive statistics	42
2.4 Conditional success rate and choice of curriculum	44
3 The aspiration level	46
4 Self-confidence	50

4.1	Estimation bias	51
4.2	Miscalibration bias	51
4.3	Synthesis on confidence	53
4.4	Aspiration and the miscalibration parameter	55
5	The decision model	56
5.1	Effort and estimation bias	58
5.2	Effort and miscalibration	58
5.3	Effort and aspiration levels	58
6	Estimation of the model	59
6.1	Estimation of the miscalibration parameter $1 - \mu$	60
6.2	Aspiration upgrade and its effect on success	62
7	Discussion	68
8	Conclusion	70
A	Appendix	72
A.1	Tables	72
A.2	Figures	72
A.3	Proofs	73
A.4	Instructions of the experiment (translated from French)	74
A.5	Comprehension questionnaire for a GT subject in the A-B order	82
2	Gender Differences, Feedback and Competition	85
1	Introduction	86
2	Experimental design	89
3	Research questions and description of methods	95
4	Results	97
4.1	Changes in performance and confidence assessments	97
4.2	Difference-in-difference analysis	101
4.3	Regressions	107

4.4	Welfare analysis	113
5	Discussion and conclusion	117
B	Appendix	119
B.1	Tables	119
B.2	Instructions	120
II	Creativity, Personality and Teenagers' Behaviors	125
3	Creativity, Personality and Schooling Achievement on Field Data	127
1	Introduction	128
2	The data	133
2.1	Measuring creativity using the Evaluation of Potential Creativity (EPoC)	133
2.2	The BB5 questionnaire	137
2.3	Inconsistency based on a risk aversion measure	138
2.4	Grades collected	139
3	Results	142
3.1	Who are the creative pupils?	143
3.2	Estimation of schooling achievement	145
3.3	Estimation of the probability of passing the middle school final exam	151
4	Discussion	153
5	Conclusion	154
C	Appendix	157
C.1	Figures	157
C.2	Tables	158

4	A Holt & Laury Measure of Risk Aversion on Teenagers	163
1	Introduction	164
2	The data	169
2.1	Risk Aversion Measure	170
2.2	The BB5 questionnaire and demographics	171
2.3	Measuring creativity via the EPoC procedure	173
2.4	Grades collected	175
3	Attitude toward risk	175
4	Explaining inconsistency	181
5	Discussion and conclusion	185
D	Appendix	188
	Appendix of Part II	189
	General Conclusion	205
	Bibliography	227
	List of Tables	230
	List of Figures	232

Résumé

Cette thèse apporte une contribution expérimentale au rôle des facteurs psychologiques dans des choix éducatifs impliquant prise de risque et performance.

L'utilisation des théories psychologiques dans l'étude des comportements économiques est devenue une discipline à part entière. Le but principal consiste à expliquer les mécanismes impliqués dans la prise de décision. La raison principale provient du fait que les modèles économiques "classiques" sont souvent remis en questions dans leur manière d'appréhender le processus de décision. Bien que ces modèles constituent naturellement une première étape pour l'élaboration des politiques économiques, les hypothèses sur lesquelles ils reposent peuvent être une représentation simplifiée du comportement de l'agent.

Nous nous limitons dans cette thèse à l'analyse de l'impact de certaines variables psychologiques: les préférences pour le risque, la confiance, la personnalité, la créativité. Nous évaluons également l'effet du genre. Les décisions et les comportements étudiés se réfèrent systématiquement à des choix éducatifs, bien que la plupart des chapitres se situent dans un contexte plus général et peuvent donc s'appliquer à d'autres domaines impliquant les mêmes mécanismes tels que les choix de carrière.

La première partie de cette thèse évalue le rôle joué par la confiance (en ses propres capacités) dans la prise de décision lors de deux types d'activités risquées: la décision (individuelle) de continuer ou non l'activité impliquant un accroissement de sa difficulté, et la décision d'entrer en compétition. Ces deux activités reposent sur une distribution des probabilités de succès inconnues, la confiance étant alors la probabilité subjective de réussite. Dans un cas (chapitre 1), la confiance en soi est absolue et peut varier avec le niveau d'aspiration et le niveau des capacités. Dans un autre cas (chapitre 2),

la confiance en soi est relative et peut être modulée par la réception d'un feedback et selon le genre de l'individu. Nous montrons que la confiance en soi, qui varie avec ces différentes dimensions, a un impact sur les performances et les décisions observées.

Dans la deuxième partie, nous nous intéressons à la population des adolescents. Deux chapitres étudient l'influence de variables psychologiques sur la réussite scolaire et les préférences pour le risque. Dans le chapitre 3, une étude de terrain, nous tentons d'expliquer la réussite scolaire (mesurée via les notes obtenues pendant l'année) par la personnalité et la créativité. Nous observons aussi un effet genre différencié sur la réussite en classe et au brevet (examen national et anonyme). Le chapitre 4 vise à améliorer la compréhension des préférences pour le risque de l'adolescent afin d'optimiser les politiques qui leur sont destinées.

Mots clés: Déterminants psychologiques, décision, éducation, risque, performance.

Abstract

This thesis contributes to the experimental literature on the role played by psychological factors in educational choices implying risk-taking and performance.

Using psychological insights to study economic behaviors has become a discipline on its own. Its main goal is to explain most of the mechanisms involved in the process of decision making process. The reason for introducing concepts of psychology into economic theory is mainly because "classic" economic models are often questioned on the decision-making process they propose. Even though these models are a first step for the elaboration of economic policies, the assumptions made are usually a simplification of the economic agent's behaviors.

This dissertation considers an analysis on specific psychological determinants: risk preferences, self-confidence, personality and creativity. We also evaluate gender effects. The studied decisions and behaviors are systematically related to educational choices, even though most of the chapters are presented in a more general framework and can thus be applied to other decisions implying the same mechanisms such as career choices.

The first part is dedicated to investigate the effect of confidence (in one's own abilities) on two types of decisions that involves a risky dimension : the (individual) decision of continuing the activity or not with an increasing difficulty, and the decision to enter competition. Both of these activities rely on an unknown distribution of probabilities of success. Confidence therefore becomes the subjective probability. First, (chapter 1), we consider absolute self-confidence which can vary with aspiration and ability levels. Second, (chapter 2), we focus on relative self-confidence which can be modulated by a feedback receipt and the gender of the individual. We show that self-confidence, which varies with these different dimensions, has an impact on performances

and observed decisions.

The second part of this dissertation focuses on teenagers. We present two studies on the psychological determinants of schooling achievement and on risk preferences. Chapter 3 examines if schooling achievement (measured by schooling grades during the year) can be explained by creativity and personality. We also observe a different gender effect on a national exam success and schooling grades. Chapter 4 improves the understanding of teenagers' risk preferences in order to improve policies aimed for them.

Keywords: Psychological determinants, decision, education, risk, performance.

GENERAL INTRODUCTION

When Economics meets Psychology: a study on the psychological determinants of risk-taking, education and economic performance.

Using psychological insights to study economic behaviors has become over the last 60 years a discipline on its own. Its main goal is to explain most of the mechanisms involved in the process of decision making. The reason for introducing concepts of psychology into economic theory is mainly because economic models are often questioned on their ideal process of decision-making (Rabin, 1998). Even though these models are a first step for the elaboration of economic policies, the assumptions made are usually a simplification of the economic agent's behaviors. Consequently, it has been shown in many researches that economics can greatly benefit from psychology theories and methods.

Simon (1986) wrote *"Everyone agrees that people have reasons for what they do. They have motivations and they use reason (well or badly) to respond to these motivations and reach their goals."* The idea here is not to question the rationality of the individuals *per se*, but to understand some of the processes of economic decision-making by using several psychological cues. This field raises various questions. How can we introduce psychological variables in economic models? Can we generalize an observed behavior? What happens when we digress from the standard economic models?

A decision is based on considering the different states one evolves in, the probability of occurrence of each event belonging to each state, but it is also determined by the environment¹ and the agent's characteristics. Among these characteristics we can hold their preferences (towards time, risk or social preferences for example) that are revealed by their observed choices and actions, and their psychological states that can rely on

1. By environment we mean the social or cultural environment for instance.

their emotions, self-confidence, motivation, creativity², and intrinsic variables (such as personality or gender) etc... All of these components interact during the decision process to result in a unique outcome. The difficulty is then to disentangle their respective role in the creation process of this outcome and in various economic behaviors in order to improve models' predictions. As economic decisions can be applied to infinite situations for all types of economic agents (firms, consumers...), this dissertation limits its study to specific decisions that involve risk-taking activities and that can be related to educational or career choices. Hence, even though some of its research applies to a more general setting, discussions always try to relate the results to these two fields.

This thesis presents two main parts, respectively organized in two chapters, that study some of the determinants implied in the decision process that are: confidence, risk preferences, personality, gender and creativity.

The first part is dedicated to analyze the effect of confidence on two types of decisions that involve a risky dimension and that can be applied to the educational domain: the decision to undertake an individual risk-taking activity, and a tournament-entry decision, in which performance determines the final outcome of success or failure. Risk-taking can be evaluated by pure risk preferences but also by observing the individual's actual decision. Performance is determined by the ability level but also by the confidence and the environment's nature (a competitive environment could, for example, represent more pressure).

Even though decisions are made in an experimental context-free setting, the underlying decisions we hence try to replicate are the following:

- Once one reaches a level of education, the choice of continuing to a higher level or to stop can arise. The decision is risky in the sense that continuing implies a risk of failing that depends on the ability, but more directly on the actual

2. As a factor rarely considered.

performance. In case of success, the agent is better-off than if she had stopped, as one additional year of study (or a higher diploma) should generate a higher wage on the labor market. But in case of failure, the outcome becomes smaller than if the agent had stopped one level before (one more year of study is costly in time but also financially and there is no return in this latter case³). Probabilities of success are however unknown as success depends on the absolute performance. It is then confidence in one's own absolute abilities that will take part in the decision process. The question is to find out to what extent.

- Educational choices can also be tournament-entry decisions. Some educational tracks imply a competitive environment: taking or not a competitive exam to get into a top university which is associated with a higher wage on the labor market than a less selective university⁴. It is a risky decision because; on one hand, if the agent fails the competition, she cannot continue into the selective track and has borne the psychological and/or monetary cost of undertaking the test; on the other hand, if she had chosen a surer but less selective track, the outcome would have been smaller than the one after the competition in case of success, but greater than in case of failure. As far as competitive behaviors are concerned, the performance of the other competitors matters and determines one's success. It is then relative confidence (that is, the confidence one has in one's own abilities compared to others') that will come into play.

Hence, in these two decisions, the determinants of risk preferences, confidence and performance play a big role. These chapters contribute to the literature on confidence and decision-making by going further into the analysis of confidence: we show that it

3. Other examples can be considered: the consequences of a failed diploma or an interview with the employer, entering the labor market later than expected etc..

4. In France, the system of "Grandes Ecoles" is typically an illustration of this case. The student has the choice between a regular university, less selective and a "grande école" in which one can enter only after succeeding a difficult competitive exam.

can vary with aspirations and ability levels (in the first chapter), and with feedback receipt and gender (in the second chapter), and that this can modulate choices.

The second part of this dissertation focuses on the population of teenagers and presents two studies on the psychological determinants of schooling achievement and on risk aversion.

First, as this thesis is concerned with the determinants of performance, we investigate some of the variables that can impact schooling grades. The latter are imperfect indicators of true abilities⁵ but they are the main determinants of future schooling and orientation choices. We consider here how personality, creativity and gender impact schooling grades. Another implicit question is to evaluate is enhanced or appreciated in the French educational system.

Second, it is sometimes difficult to rightly set up policies targeting children or young adults because of a lack of knowledge of their preferences and behaviors. We choose to focus on their risk preferences and present a methodological work to elicit them in order to better understand what shapes them. We propose to implement an experimental economics elicitation procedure on a sample of teenagers and explain why having a better knowledge of these preferences can improve policy setting and implementation. This part contributes to the literature interested in the effect of certain types of abilities (non-cognitive vs. cognitive) with an emphasis here on creativity and to the literature which focuses on young agents' preferences.

The first section of this general introduction will describe the methods used in the different chapters of the dissertation. Four sections are then devoted to describe the different determinants and their measurement (if it is suitable) and refer them to the corresponding chapters. The last section of this introduction presents an outline of the

5. Even though a pupil can have high ability in math, he can get a bad grade at a math test because, among other possible explanations, he is not conscientious enough, or dissipated or because the teacher unconsciously favors the others

thesis, summarizing each of the four chapters.

0.1. The empirical method

The work presented in this thesis studies specific psychological determinants involved in economic decision-making and largely adopts an empirical approach. It is based on economic theories of decision and econometric analysis. While the first two chapters use laboratory experimental data, the last two are based on data gathered in the field.

Experimental economics in the laboratory has been inspired by experimental psychology and has added specific features of its own such as monetary incentives, context-free and no deception (see [Croson \(2005\)](#), [Hertwig and Ortmann \(2001\)](#)). The first goal of this method is to test predictions of theoretical models. The main advantage of laboratory experiments is that we are able to isolate the specific variables we are interested in while controlling for the environment. Moreover, this method allows to dispose of some econometric problems coming from missing variables or sample sizes ([Montmarquette, 2008](#)). One may however worry about its external validity⁶ and on the sample composition because this latter is usually made of graduate students willing to participate in the experiments. This raises the question of whether the observed behaviors or decisions in the lab on a sample of graduate students can be generalized to any other population.

Targeted populations can in fact be reached more easily with field data. Going in the field implies meeting the sample in its environment which improves the external validity of the results. Nevertheless, as with all panel survey dataset, we may face problems of selection bias and omitted variables. Some controls remain difficult to implement in the field as there are specific time, technical and ethical constraints, on young individuals.

6. The question being if experimental results can be valid outside of the lab.

We believe that the experimental method is appropriate when studying a specific decision that can be noisy outside of the laboratory that is where it becomes impossible to control for variables that cannot be observed. Field data are quite powerful when we need to control for the environment and when we need to reach a specific population.

More precisely, chapters 1 and 2 present controlled laboratory experiments based on real-effort tasks where decision-making is observed and analyzed. Chapter 3 and 4 are based on the same type of data collected in the field. However, while chapter 3 provides more of an empirical study, chapter 4 relies on an experimental method (done in the field).

The following sections present the different psychological determinants studied in this dissertation by defining them and placing them within the perspective of analyzing economic behaviors.

0.2. Risk preferences

Economic theories represent risk as a probability on the outcome of any action and decision an agent may make. The economic agent introduces this probability in her decision process and according to her tolerance to risk, she makes a more or less risky decision. Introducing risk in choice-theoretic models as well as having a good measure of risk preferences is then crucial in order to understand economic behaviors that could help defining better policies which involve risks.

Many economic decisions involve risk (known probability distribution) and uncertainty (unknown probability distribution)⁷. Its predominance in life explains why economists have been interested in explaining and understanding the process of decision-making under potential risks.

7. Uncertainty arises when one is uncertain about the possible outcome of an event or one cannot assign probabilities to the possible outcomes. For instance, in the case of an environmental catastrophe as an earthquake, one is not able to know the exact probability of occurrence as well as the exact magnitude of the outcomes.

One of the first models on risk which introduces probabilities into the decision process is the Expected Utility (EU) Model (Von Neumann and Morgenstern, 1945) where agents are assumed to maximize their expected utility according to the probability of the different outcomes they may face. Theoretical works extended this vision by introducing psychological insights on the perception of probabilities. In fact, under the EU theory, decision makers consider probabilities linearly and no distortion may be encountered. Subjective probabilities was later introduced into the Subjective Expected Utility (SEU) model of Savage (1954). However, two rational agents, facing the same situation and having the same information, may develop different beliefs and preferences. Evidence, especially experimental (Ellsberg, 1961, Slovic and Lichtenstein, 1968, Kahneman and Tversky, 1979), shows that people do not weight probabilities in a linear manner and frequently violate the Savage's independence axiom (cf. Allais (1953) paradox).

Hence, theories such as Prospect Theory (PT) (Kahneman and Tversky, 1979, Tversky and Kahneman, 1992) and the Rank Dependent Utility (RDU) models (Quiggin, 1982) are various models taking into account these different elements of discussion.

From an empirical point of view, in the case of facing a risky activity, the researcher has to find a way to distinguish the individual perception of risk from the other decisional variables. In order to elicit risk preferences, economists prefer lottery-based questionnaires with real incentives or treatment tasks such as:

- Price list design where subjects are given a list of binary lottery choices to make all at once (widely used in experimental economics).
- Random lottery pairs where subjects face binary lottery choices in a sequence and must choose the preferred lottery.
- Ordered lottery selection: the subject chooses one lottery from an ordered set.

- Becker-DeGroot-Marschak (BDM) design (Becker, DeGroot and Marschak, 1964): "individuals are asked to state their minimum selling price for a risky lottery. The experimenter then draws a random number between the lowest and the highest outcome of the lottery. If the price that the individual states is lower than or equal to the drawn number, she receives the drawn number as her payoff. Otherwise she has to play the risky lottery" (quotation from Blavatsky and Köhler (2007)).

Harrison and Rutström (2008) review the different risk elicitation methods and list their advantages and disadvantages.

Psychologists consider risk taking as a complete intentional and repeated commitment into a dangerous decision (alcohol consumption, drugs, motorcycle driving etc.). It can depend on one's own risk perception and on individual characteristics (sensation seeking, impulsivity...). Nevertheless, they use different types of measure to elicit risk-taking propensity. One of them is the self-report questionnaires measuring personality traits related to risk-taking (see Rohmann (2002) for a review on the different tests): sensation seeking (Zuckerman, Eysenck and Eysenck, 1978), venturesomeness (Eysenck et al., 1985), impulsiveness (Barratt, 1985, Eysenck et al., 1985) or deficit in behavioral constraints (Tellegen, 1982). Questionnaires target different domains in which risk behaviors may arise, such as financial, social, sport, health... Since the 90's, in order to overcome some shortfalls from these tests (mainly because of their declarative nature), indirect measures have been elaborated and combined with direct measures: the Objective Personality Test. We can cite as an example the *Balloon Analogue Risk Task* (Lejuez et al., 2002) where participants are asked to blow up a balloon, which increases the payoffs, until it bursts. It has been found that in teenagers, this score task is correlated with risky behaviors such as consuming drugs or delinquency (Aklin et al., 2005). Another type of measure is the *Game of Dice Task* (Brand et al., 2005) where participants roll a die many times and bet on the number that will appear. One

can bet on 1 to 4 numbers associated with different gains and losses.

Regarding this thesis, we do not specifically test or question theories but rather focus on how the perception of risk may influence decisions. The four chapters address risky decisions in different manners. As long as a decision has to be made with different possible outcomes which depend on their probability of occurrence, risk preferences play a major role. The question is then how do we measure risk preferences and how do differences in these preferences imply various choices among individuals?

In the first two chapters of this dissertation, which deal with two types of decisions where risk is involved, preferences are not directly measured. The role played by risk preferences is elicited through the variation of the context between tasks and the observed decisions.

In chapter 1, risk aversion is mainly considered theoretically in a model showing how risk preferences modulates the level of effort provided to succeed in a task and impact then the decision to continue.

Chapter 2 deals with the decision to enter a tournament. When deciding to enter a competition, risk preference may interact with the pure taste for tournament. The risk encountered in this situation is to lose the tournament and thus earn nothing, whereas the alternative of piece-rate wage-based becomes a surer option as one is certain to earn a payoff. In order to distinguish the only will of entering a tournament from risk aversion, we use a specific succession of tasks firstly implemented by [Niederle and Vesterlund \(2007\)](#) and followed by many others, that is carefully described.

Risk preferences are carefully measured and take a central attention in the last two chapters (3 and 4). Risk aversion was elicited on teenagers via the procedure of [Holt and Laury \(2002\)](#) (price list design) where pupils from a middle school had to make ten choices between a sure lottery vs. a risky lottery. The degree of risk aversion is directly observed by looking at the number of sure lotteries a subject has chosen.

0.3. Confidence

From a standard theoretical perspective, confidence is a distorted probability of success. It is a subjective belief most of the agents have and that can be modified across time with new information (like a series of failure and/or success, see for example [Compte and Postlewaite \(2004\)](#)) which leads to an update of one's beliefs. In fact, according to the available information about themselves and about the environment, individuals will update these beliefs in order to improve their decisions. Along this dissertation, we only consider self-confidence in one's own success and abilities (and not in the probability of occurrence of a self-irrelevant event).

As mentioned before, experimental evidence shows that objective probabilities can be transformed by the agent into subjective probabilities which can represent the confidence one has for any event to occur. Psychology and economics agree on the fact that the psychological state of mind of an individual can affect decisions and performance. Stress, fear, anger, confidence and many other emotions participate in modulating the psychological state and thus indirectly affect economic decisions and performance.

Indeed, self-confidence can be a motor for performance by stimulating effort and the will to succeed. It has been shown, however, that over-confidence can have a negative impact on decision leading to non-optimal decisions, as well as on success, especially when over-confidence is based on abilities ([Bandura, 1993](#)). The same problems may arise with under-confidence. Consequently, differences in levels of confidence among individuals will imply different decisions.

When addressing the question of rationality of beliefs, economists would traditionally use the Bayesian approach. Bayesian beliefs rely on prior probabilistic beliefs on any uncertain facts or events. When new information occurs, the Bayesian prior will be updated by the Bayes' rule. As the agent has to make a decision, she maximizes her expected utility weighted by her updated Bayesian beliefs. However, the Bayesian

approach has been criticized by decision theorists and experimentalists. This reasoning first lacks identifying which beliefs are rational meaning that subjective beliefs can be rational, and second, how they are generated, that is looking for the determinants leading to these beliefs (Tversky and Kahneman, 1974, Gilboa, Postlewaite and Schmeidler, 2009). Recent theories incorporated the formation of Bayesian beliefs into decision models (Compte and Postlewaite, 2004, Gilboa, Postlewaite and Schmeidler, 2009). Experimentalists focus on testing the assumptions of Bayesian rational beliefs update and try to understand how the formation of the beliefs will impact decisions and performance (Möbius et al., 2011, Ertac, 2011).

The experimental analysis of confidence sheds light on specific and individual characteristics-related variations (Stankov et al., 2012). Many papers show that confidence depends on cognitive abilities (which reflect true abilities), on non-cognitive abilities, but also on gender, age, cultural differences or social background. Hence, these individual characteristics will impact beliefs or confidence in a different way leading to a different decision-making process.

One important characteristic of confidence⁸ is that it is endogenous. It can be updated with new incoming information such as feedbacks about oneself or the others⁹ or the environment. It has a crucial role in many educational and economic situations that directly relate to performances such as schooling choices and success, and on the labor market outcomes (see for instance Heckman (2000), Möbius and Rosenblat (2006)).

Confidence elicitation is a self-declared measure that can be either uttered in an absolute manner ("State the probability between 0 and 100 you think you are right") or in a relative manner ("Tell if you think you are relatively better than the others").

8. Psychologists rather use the concept of self-efficacy (Lent, Brown and Larkin, 1986, Bandura, 1993)

9. We actually implement a performance feedback in chapter 2 and see its effect on confidence.

The difference between psychologists and economists, is that the latter incentivize confidence elicitation such that they pay subjects according to the accuracy of stated beliefs in addition to payments for other decisions¹⁰ whereas psychologists do not¹¹.

It is generally observed that people are not accurate about their success usually leading to what we call a confidence bias. This bias corresponds to the difference between the confidence rate and the true probability of success. This bias measures the calibration degree. Research shows that most of the people are not well calibrated, and that confidence bias, especially over-confidence, has an effect on decision-making (see [Lichtenstein, Fischhoff and Phillips \(1982\)](#) for a review). The difficulty of the task can affect overconfidence. People are known to be all the more confident in their relative ability that the task is easy¹².

Most of the time individuals tend to be overconfident whatever the difficulty task. Overconfidence arises when the difference between one's own subjective belief of success and the objective (true) probability of success is positive. In a review, [Moore and Healy \(2008\)](#) identified three distinct types of overconfidence: overestimation, overplacement and excessive precision. The overestimation is the most common in the research literature; it is related to estimating too highly one's actual ability or performance. Overplacement occurs when an individual believes that he or she is better than the others; for instance in a sample of American drivers, 93% of subjects reported themselves as more skillful and less risky than the average driver in USA ([Svenson, 1981](#), [Benoît and Dubra, 2011](#)). Finally, the overprecision occurs when people are too

10. Various methods exist. The most used is the Quadratic Scoring Rule (QSR): the agent reports the discrete probability distribution $p = (p_1, \dots, p_n)$ where p_i represents the reported probability that event i occurs. If the event j occurs, the QSR yields a payoff equal to $Q_j(p) = a + 2b * p_j - b \sum_{i=1}^n (p_i)^2$. A simpler belief elicitation mechanism is only to reward if the beliefs are correct (especially when the reported belief is not a probability but a rank for example). We use another mechanism in chapter 2.

11. Experimental psychologists display also different types of confidence assessment tests: after the administration of a test item (cognitive test), the subject is asked to give an answer and just after she has to indicate on a scale (usually a Likert-type scale) how confident she is that her answer is correct.

12. see the Hard-Easy effect and the Dunning Krueger effect described in chapter 1 ([Lichtenstein and Fischhoff, 1977](#), [Kruger and Dunning, 1999](#)).

sure that they know the correct answer.

The first chapter elaborates a model that tries to incorporate two types of biases (an estimation bias, related to Bayesian update, and a cognitive bias related to over or under estimation of one's own ability) into one unique function of self-confidence. It tests if one can disentangle both of these biases with experimental data. Confidence relies on the absolute self-confidence the agent has on succeeding a task. It is asked *ex ante* and aims at predicting future performance. However, we do not claim any causal effect from confidence to success. It might be also the other way around as the specific task is also performed before declaring self-confidence. It is then rather a dynamic process between confidence and success. Chapter 2 presents confidence in a different manner. Subjects still have to perform a task, but they are now asked to declare relative beliefs among the four possible performance quartiles. As competition choices are at stake, we are not interested in absolute but in relative self-confidence translated into beliefs of being better or worse than one's opponents.

0.4. Intrinsic Variables

Confidence and risk aversion are endogenous characteristics because they can be partly influenced. They can vary with experience, time, with a change in probabilities etc... However, there are characteristics which are very difficult to influence or cannot be changed; this is the case for personality and gender. These variables are not included in standard economic models. However, many empirical works have shown that gender or personality has an impact and in that case it is rare to find a persuasive explanation of it.

0.4.1. *Personality traits*

In psychology, the personality traits analysis is used to describe and explain behaviors. They are part of the determinants that will induce the action or the choice of a

person (Cattell, 1979).

Psychologists' works led to a specific description of personality based on five traits and commonly known as the Big Five. These five traits are Agreeableness, Conscientiousness, Extraversion, Emotional Stability and Openness. Agreeableness refers to someone who is helpful, sympathetic, and cooperative. Someone who is characterized by conscientiousness can be assimilated to someone who is disciplined, organized, and achievement-oriented. Extraversion is revealed through a higher degree of sociability and talkativeness. The emotional stability refers to someone who can control his impulse and his anxiety. And finally, openness reflects the intellectual curiosity and the preference for the novelty and the variety. A way to measure it with the Big Five test (McCrae and Costa, 1987) is described in chapter 3 and 4.

Economics recently started to have an interest in personality as a predictor of observed economic behavior. It especially concerns the literature confronting the cognitive abilities with the non-cognitive ones and their impact on economic outcomes (Bowles and Gintis, 1975, Bowles, Gintis and Osborne, 2001, Heckman, 2006, Heckman, Stixrud and Urzua, 2006). Non-cognitive skills can be academically and occupationally relevant skills and traits but that are not specifically intellectual or analytical in nature such as motivation, self-control, personality or self-confidence. Almlund et al. (2011) present an extensive review of the existing relationship between economics and personality. Among all if the considered economic topics, they present results from other studies on the relationship between personality and education attainment or labor outcomes. For example, they point out the fact that conscientiousness and openness are positively correlated with the number of years of study (Goldberg et al., 1998, Van Eijck and de Graaf, 2004) but also with job performance (Salgado, 1997, Hogan and Holland, 2003, Nyhus and Pons, 2005), whereas emotional stability, agreeableness and extraversion are negatively correlated with the number of years of study. There is also some evidence that emotional stability predicts wages (Judge and Hurst, 2007, Drago, 2011,

Duncan and Dunifon, 2012). Other personality traits such as locus of control and self-esteem have an impact on these outcomes but they will not be considered in this thesis.

In the core of this dissertation, personality traits are assumed to be rather stable over time and are used as explanatory variables to predict schooling grades (see chapter 3) and risk preference (see chapter 4) of 14-15 years old teenagers. It will also be related to creativity potential.

0.4.2. *Gender*

There is a growing literature on gender that arose in economics stemming from the fact that gender differences are often observed in empirical studies. Men and women seem to behave differently and differ in their decision process. Specifically in experimental economics, gender differences were found for risk aversion, competitive taste, social preferences and trust. In their review, Croson and Gneezy (2009) explain that in most of the experiments using probability lotteries, women are found to be more risk averse than men. However, they specify studies where this tendency changes according to cultural determinants and payoffs framing¹³. Different explanations arise for this finding but are still hypothetical. It can be the case that emotions generated by risky situations differ among gender implying diverse decisions (Brody, 1993, Loewenstein et al., 2001). In psychology as well, many studies showed that women are less prone to make risky decisions (Wagner, 2001, Hirschberger et al., 2002, Zuckerman and Kuhlman, 2000). Weber, Blais and Betz (2002) indicate that women are more risk averse in the health domain, leisure and ethics. Kerr and Vlaminkx (1997) find that women consider risky situations as more stressful compared to what men declare.

In terms of confidence, men tend to be more confident than women and this can also be a cause of a higher taste for risk-taking. Research in economics got interested in this

13. Table 1 of their paper reviews this literature.

phenomenon since it has consequences on the labor market ("gender wage gap", payment schemes choices), in educational choices and in investment decisions. Barber and Odean (2001) find that male investors are more likely to trade than female investors, assumedly because of their overconfidence in their ability. Bengtsson, Persson and Wilenbäck (2005) use natural experimental data of an Economics first year exam to assess whether male students would take an extra question, allowing a higher mark. This opportunity was more chosen by male than female students even though this difference was less accountable for older students. The gender difference on overconfidence was investigated in children's behaviors since it is believed to emerge in early life (Sutter and Rützler, 2010). Recently, an article by Dahlbom et al. (2011) use answers from 14 years old high school pupils on their expected grades in mathematics and found that boys were overconfident while girls were underconfident. This is not totally in line with previous results since they find that instead of being overconfident but less than boys, girls are underconfident. It also has been shown that the type of task and its environment could impact overconfidence. For instance, Nekby, Thoursie and Vahtrik (2008) find that women who self-select into a male-dominated environment may be at least as overconfident as men. The statement according to which women are less overconfident has been reassessed and slightly divergent results have been found. As far as tournament entry is concerned, men enter competition more easily although their performance is not necessarily significantly higher than women's (Gneezy, Niederle and Rustichini, 2003, Niederle and Vesterlund, 2007, Gneezy, Leonard and List, 2008, Datta Gupta, Poulsen and Villeval, 2012).

Apart from behavioral studies, the empirical literature has focused on the gender issue on labor or educational outcomes. Women stay nowadays longer in schools and have a lower probability of dropping out, however they still earn on average less than men all things being equal. In schools, girls often outperform boys in most of the subjects but stereotype threat leads to stereotyped tracks (*e.g.* girl students choosing more often

humanity sciences tracks, whereas boys choose more often scientific tracks). Other reasons are also investigated: teachers' grading, environmental causes or students' motivation (Lavy, 2008, Falch and Naper, 2013, Ouazad and Page, 2013).

The gender issue is studied in this thesis regarding risk-taking, taste for competition and schooling grades. Social preferences such as inequality aversion, reciprocity or altruism are not studied here even though a large literature focusing on gender issues investigates these topics. Chapter 2 specifically studies the gender gap in competitive entry. The last two chapters, even though it is not their primary goals, will consider at one point gender differences. The question of gender is thus a recurrent concern all along this thesis. For example, in the analysis of risk aversion in Chapter 4, we find that teenage girls are significantly more risk averse. As schooling grades are at stake in chapter 3, we find that girls succeed better in 9th grade but have a lower probability than boys to pass a national exam.

0.5. Creativity Potential

Creativity is a notion that will be covered in the last two chapters of the thesis in order to first, evaluate its role in schooling achievement, and second, to see whether it is related to risk preferences.

The particularity of creative potential is that it is at the frontier of the cognitive and non-cognitive abilities. On one hand, it is an ability to produce original and new ideas under certain constraints of a given situation. On the other hand, an individual who will show creativity will be considered as a creative person not in terms of skills but rather in terms of personality traits.

There are two different approaches to creativity. Some creativity research is interested in the eminent creativity to better understand the creative "genius" assumed to be possessed by a few people (like eminent artists) and how it can be transmitted. It is called *the Big C*. Other researchers would rather study the every day creativity,

the little c (Richards, 1990). In these studies, mainly psychologists are interested in creative activities in which the average person may participate each day. According to Kaufman and Beghetto (2009), four levels of creativity could be considered: (1) the mini-c which corresponds to a transforming learning (a student understands a new concept in mathematics for example); (2) the little-c which corresponds to the everyday creativity (to invent a new recipe); (3) the Pro-C, which corresponds to the creativity in a domain in which the individual is an expert; (4) the Big-C, which corresponds to an eminent creative contribution (Keynes or Schumpeter in economics, Picasso in arts).

Models of the creative process have distinguished phases of processing and types of thinking involved. One broad distinction opposes divergent thinking in which the goal is to explore multiple cognitive paths, and convergent thinking which seeks to focus on a single, perhaps optimal path. Complex creative performance tasks certainly involve both kinds of processes, in various degrees and in specific sequences that favor the generation of new ideas. In this vein, Lubart and Guignard (2004) propose that the moderate correlations observed between different creative performance tasks stem from the fact that there is a mix of cognitive operations and knowledge involved in each creative domain and task.

Several recent studies look at the link between personality (evaluated with the Big Five) and creativity showing that openness is probably the personality trait that is consistently related to creativity, and, to a lesser extent, extraversion (Feist, 1998, 2010, Batey, Furnham and Safiullina, 2010, Batey, Chamorro-Premuzic and Furnham, 2010, Dollinger, Urban and James, 2010).

We study creativity in the context of schooling because we think that the school has a role to play in helping children to develop their creative skills. Creativity is by nature valued in the arts which are creative by definition. However, in subjects traditionally taught at school as French or mathematics, although creativity can have

its place especially through the learning methods (problem definition, formulation of hypotheses, selection of relevant information, connecting information with each other), the development of creative skills may not be maximal. The French educational system receives reproach for its conformism-rewarded rather than for its creativity learning. Creativity is a key ability in the labor market, especially nowadays because one needs to be more flexible and because it allows individuals to develop other related abilities such as self-management, problem solving etc...

Chapter 3 actually questions and evaluates the role of creativity in the different taught subjects as well as specifying creative pupils. Chapter 4 presents a different approach on the relationship that may exist between creativity and risk preferences. We support the assumption that creativity is associated with risk-taking, especially in the domain of innovations and that this risk seeking should be formed and detected as early as during adolescence. This type of risk taking is beneficial to the individuals (as opposed to risky behaviors teenagers may encounter as well such as smoking, drinking, early pregnancy etc...).

0.6. Outline of the dissertation

This thesis revolves around the different determinants described in the previous sections. The first two chapters form a first part dedicated to the relationship between self-confidence - or beliefs - and different economic risk-taking activities. Both chapters are written in a general framework of choices and performance that can be applied to educational or career decisions. The second part focuses on the role of personality and creativity on teenagers' behaviors. This section will summarize the four chapters and present the main results.

Chapter 1 is a joint work with Louis Lévy-Garboua and Claude Montmarquette. It is based on a lab-experiment where we simulate a decision process with three suc-

cessive levels of increasing difficulty and choices to continue. The experiment is set so we can test a dynamic model that incorporates self-confidence and aspirations into an economic model of repeated choices and effort under risk. Subjects face an increasingly difficult task whose payoff is exclusively dependent on the realized performance. We assume that agent have an imperfect knowledge of their own ability, whereby both their estimation of performance ex ante and their beliefs may be biased. In our model, Bayesian agents present a self-confidence¹⁴ level function of their ability-estimations bias based on the perception of their ability (estimation bias) and of a cognitive bias leading to over-confidence for low-ability agents, and under-confidence for high-ability agents (miscalibration bias)¹⁵. We introduce as well aspirations as the goal an individual wants to reach or the level she intends to perform. It can be no lower than the status quo and may shift after each achievement and is thus composed of an endogenous and and exogenous part.

We show that a positive ability-estimation enhances effort if ability and effort are complementary factors of success (and the reverse if they are substitutes), but that miscalibration bias increasing self-confidence will reduce effort and success. We also show that for risk averse subjects, aspiration level has little effect on low ability subjects' effort but will increase high ability subjects'.

By a simple manipulation of endowments, we experimentally simulated three levels of aspiration, that we matched with the three levels of difficulty. We show that confidence and success do not interact the same way with aspirations and effort across the two ability categories of agents and we confirm the physiological assumption. A new result emerges from our analysis: we show that confidence has a good and bad component meaning that it does not always predict success. We also confirm the aspiration

14. Self-confidence is defined as the subjective probability of success which may deviate from the true probability of success.

15. We make a robust assumption coming from the psychology literature that relies on two main effect: i) the hard-easy effect (Lichtenstein and Fischhoff, 1977, Griffin and Tversky, 1992) and ii) the Dunning-Kruger effect (Kruger and Dunning, 1999).

upgrade after a positive update of abilities. A discussion is further developed on the different possible applications (notably on education).

The second chapter (2) is a joint work with Marie-Pierre Dagnies. It analyses one's reaction after providing feedback on relative performance in terms of competitive entry. It is articulated around two main research questions. One is to evaluate how participants update their beliefs after getting the information on their relative performance. The second one is to specifically study how men and women react to this information in terms of the decision to enter competition.

Our first result is that subjects update their beliefs following performance feedback more than would do a Bayesian agent. Both men and women are more pessimistic than a Bayesian agent following negative feedback; we find the opposite effect after positive feedback. Both of these effects are stronger for women than for men. Our paper also shows that low-performing participants adapt their tournament entry decision to the ability level of the competition, while high-performing participants do not.

Men and women do not react to the feedback in terms of competitive entry in the same way. While women are especially sensitive to information on their own performance, men react more strongly to the level of their competitors. Feedback does not therefore seem to be processed in the same way by men and women. Low-performing women consider their performance level *per se* while men think there is room for improvement. Regarding the efficiency of choices, men mostly enter in about the proportion suggested by payoff maximization, but not enough women choose the tournament.

Chapter 3 (a joint work with Jean-Louis Tavani and Maud Besançon) and 4 use a different approach and are based on the same dataset. A middle school in the Parisian suburb opened its doors for us so we were able to gather data on 9th grade pupils. Although using the same dataset, each chapter has its own specific purpose.

The third chapter is in line with the recent literature on the effect of non-cognitive abilities on educational outcomes (Bowles and Gintis, 1975, Bowles, Gintis and Osborne, 2001, Heckman, 2006, Heckman, Stixrud and Urzua, 2006). It presents an exploratory study and aims to go deeper into the analysis of explaining schooling achievement. It investigates the influence of creative potential, personality and other individual characteristics on academic achievement of middle school students. We focus our analysis on a largely understudied variable in economics of education, but which might have implications, that is creativity potential. By using a test of creativity elaborated by psychologists (Lubart et al., 2003), a test of personality based on the Big Five (McCrae and Costa, 1987) and a measure of inconsistent choices as well, we are able to extract some of the determinants of better grades.

The main result of our paper is that one dimension of creativity negatively predicts subjects' grades. It seems that being creative on that specific aspect does not improve performance at school. Openness and to a lesser extent conscientiousness are the only personality traits that exhibit significant positive estimates. A inconsistency variable, coming from our risk aversion measures, is negatively associated with math and physics' grades. Overall, girls have significantly higher grades than boys, except in sports, but have a lower probability to pass a national exam.

Chapter 4 focuses on teenagers' risk preferences. The way young economic agents make their decisions may have a crucial impact on economic policies efficiency. Teenagers differs from adults in the way they make their decisions and behave. There is a growing literature to understand the children's and teenagers' preferences and how they make their choices (Eckel et al., 2012, Sutter et al., 2013, Harbaugh, Krause and Vesterlund, 2002). The procedure created by Holt & Laury (2002) is widely used in experimental economics but had not yet been used on young individuals. Hence, we decided to implement this method on a sample of teenagers, first in order to verify if it can be applied to this specific sample, and second, to analyze determinants of risk aversion at

this age.

Our main finding is that the pattern of risk aversion using H&L procedure is not very different from adults found in other papers. We observe that girls are more risk-averse than boys. This result is mainly driven by two personality traits: girls who have a higher openness score and a lower emotional stability score tend to be less risk averse compared to boys. One of the creativity dimensions measure (showing the ability of producing many different original and appropriate ideas) also decreases the propensity of being risk averse, which is in line with the idea that being more creative leads to taking more risks. Concerning inconsistency yielded by this specific measure, the rate of inconsistent subjects is relatively the same as shown in other studies (26.26%) therefore showing that teenagers are not really more inconsistent than adults. Hence, the H&L procedure can be applied on a sub-sample of teenagers. Higher grades in scientific subjects decreases the probability of being inconsistent. In addition, a higher score in two dimensions of the creativity measure based on verbal tasks decreases the probability of being inconsistent.

Part I

Confidence and Economic Behaviors

CHAPTER 1

**CONFIDENCE, ASPIRATION AND
PERFORMANCE**

This chapter is a joint work with Louis Lévy-Garboua and Claude Montmarquette.

1. Introduction

When undertaking lengthy studies or a career, individuals envision reaching upper positions on the job or education ladder if they realize a good performance. In the standard economic model, their repeated choices of education, job, and effort essentially depend on their abilities and discount rates. Risk aversion will also play a role if performance has a random component. In this model, risk and time preferences, captured by risk aversion and the discount rate, are viewed as important determinants of economic choices. However, they are assumed to be the only psychological determinants of these choices. In the present paper, we wish to introduce two additional psychological factors in the economic model of choice: self-confidence (Bénabou and Tirole, 2002, Compte and Postlewaite, 2004) and the aspiration level (Boudon, 1973). These factors, which are commonly considered in psychological and sociological studies, can be straightforwardly introduced in a Bayesian model with choices to continue or to quit a risky activity.

Our paper aims to explain the individuals' decision to continue to a higher difficulty level taking into account their ability, effort, self-confidence¹ and level of aspirations. It is based on a lab-experiment where we simulate a decision process with three successive levels of increasing difficulty and choices to continue. The experiment is set so we can test a (dynamic) model that incorporates self-confidence and aspiration level into an economic model of repeated choice and effort under uncertainty for an increasingly difficult task whose payoff is exclusively dependent on realized performance. The effort provided is then co-determined with the decision to continue the task, knowing that the difficulty will increase.

Compte and Postlewaite (2004) present a model where the agent has to decide whether to undertake a risky activity or not. They assume that emotions can affect

1. We consider here absolute self-confidence in one's own success. We do not take into account confidence of self relative to others, nor confidence of the precision of one's estimate.

performance and build a standard decision-theoretic model where emotion depends on past successes and failures. It is a Bayesian model where the probability of success depends on self-confidence: if this agent is unsure about her ability, then the probability of success decreases. They make two main assumptions: a higher confidence implies a higher probability of success, and confidence depends on the perception the agent has on her past successes. They incorporate the attribution bias where the agent explains success with internal reasons ("I did it"), and failure with external reasons ("I was not lucky"). They show that even though Bayesian agents exhibit this bias, their decisions can still enhance welfare if confidence positively affects performance.

Our model is in the spirit of [Compte and Postlewaite \(2004\)](#). However, it is more parsimonious and rests on different behavioral assumptions than theirs. A gain in parsimony is obtained by not introducing in our model confidence or optimism in the "performance technology", that is, as a direct determinant of the frequency of success. This has the advantage -in our opinion- of avoiding an unwarranted disconnection between perception, which affects confidence, and beliefs, which affect decision. Our performance and decision technologies are standard, with effort and ability as the sole factors of production, and confidence describing beliefs. On the other hand, we assume that agents have an imperfect knowledge of their own ability, whereby both their estimation of performance *ex ante* and their beliefs may be biased. We do not assume attribution bias as failure in our framework ends the game. We make a more robust assumption coming from the psychology literature that relies on two main effects: i) the hard-easy effect ([Lichtenstein and Fischhoff, 1977](#), [Griffin and Tversky, 1992](#)) and ii) the Dunning-Kruger effect ([Kruger and Dunning, 1999](#)). The hard-easy effect is based on the relationship between the perception of the difficulty of the task and ability such that people underestimate their ability to perform an easy task and overestimate their ability to perform a difficult task. This perception of one's abilities will affect confidence and thus the level of effort provided. The Dunning-Kruger effect makes a distinction

between the ability level of an individual and asserts that the low ability individuals overestimate their abilities whereas the high ability individuals underestimate their abilities, without considering the difficulty level of the task². Taking these effects into account imply a consideration of the ability level of the agent to predict her confidence and level of effort provided. It is then assumed that the individuals do not perfectly know their true level of abilities and have a biased perception of the latter. We define self-confidence as the subjective probability of success which may deviate, at least temporarily, from the true probability of success (observed). In our model, Bayesian agents present a self-confidence level function of their ability-estimation bias based on the perception of their ability (Bayesian update or estimation bias) and of a cognitive bias leading to over-confidence for low-ability agents, and under-confidence for high-ability agents (miscalibration bias).

We define aspirations (Quaglia and Cobb (1996) offer a nice review of the concept in historical perspective) as the goal an individual wants to reach or the level that she intends to perform, which can be no lower than the status quo and may shift after each achievement (Hoppe, 1976). Reaching an ambitious goal, that is, fulfilling high aspirations, is a risk-taking activity which yields a high payoff when performed successfully (success) and a low payoff when performed unsuccessfully (failure). The status quo is an alternative activity yielding a sure payoff in between these two outcomes. We introduce aspiration by considering a reference-dependant utility function and by taking the agent's aspiration level as reference, which can either be the status quo or a more ambitious goal. This specification also allows to take into consideration risk aversion and its effect on decision. Aspiration depends on two components: one exogenous which is the initial induced level of aspirations, and an endogenous component

2. Both effects rely in fact on the same idea: if a task is considered as easy, this means that one has the ability to perform it. According to the hard-easy effect as well as to the D-K effect, individuals should be then underconfident in succeeding the task. The opposite reasoning can be made for a hard task, that will be consequently performed by low ability agents who should then exhibit overconfidence according to both effects.

that represents the upgraded aspiration after a success (by updating the perception of one's own abilities). In specific situations, high aspirations will enhance confidence, effort, and performance. The model elaborates the risky decision of continuing to a higher difficulty level which depends on: the aspiration level and the confidence level³ (as a subjective probability of success incorporating the estimation bias on the ability estimation and the cognitive bias).

We test the model with our experimental data where time discounting is unlikely, so discounting behavior is not taken into account in the model (see [Bénabou and Tirole \(2002\)](#), [Gervais and Odean \(2001\)](#) for behavioral assumptions on the discount factor). During the experiment, subjects have to carry out a real-effort task with three levels of difficulty which can lead to success or failure. This allows us to distinguish low and high ability participants based on their performance in the task so we can test the behavioral assumptions of the model. High ability individuals should pass the lower level and some at least should succeed the highest level. Confidence in succeeding each level is elicited twice: during level 1 and just before level 2 (conditioning on succeeding level 1 and continuing to level 2). We simulate three initial levels of aspirations by framing the payoffs as either gains or losses relative to an initial endowment leading to three treatments: a Loss Treatment (LT) in which subjects are well-endowed with €35, an Intermediate treatment (IT) endowed with €20, and a Gain Treatment (GT) in which subjects are not endowed at all. Thus, the LT and the IT subjects are initially placed in a loss framework; the further they go on with the experiment the less they may lose. In contrast, the GT subjects can only earn money. A specificity with the IT is that subjects can experience a change in their reference point and go beyond, allowing them to update their aspirations. We hence induce an initial aspiration by framing the payoffs (treatments), but the participant can revise her aspiration level in

3. Large individual differences in aspiration and self-confidence may be caused by differences in personality ([Gardner, 1940](#)) and social origin as emphasized by sociological theories of education ([Bourdieu and Passeron, 1964](#), [Boudon, 1973](#), [Duru-Bellat, 2003](#)).

case of unexpected success (coming from the exogenous and endogenous components)⁴.

We make four assumptions:

H1 Low-ability subjects over-estimate their ability and are thus overconfident.

H2 High-ability subjects under-estimate their ability and are thus underconfident.

H3 Aspirations interact with abilities which impact confidence and success.

H4 Unexpected success can imply a modification of aspiration which can impact confidence and future achievement.

The model shows how the decision to continue to a higher difficulty levels takes into account all of these psychological variables (see [Breen and Goldthorpe \(1997\)](#) for an application to education with different levels of aspirations) and how the effort provided in the task will vary. The specification of the confidence function highlights the differentiation between the estimation bias and the miscalibration bias. We show that a positive ability-estimation enhances effort if ability and effort are complementary factors of success (and the reverse if they are substitutes), but that miscalibration bias increases self-confidence but will reduce effort and success. We also show that for risk averse subjects, aspiration level has little effect on low ability subjects' effort but will increase high ability subjects'. By using our experimental data, we show that confidence and success do not interact the same way with aspirations and effort across the two ability categories of agents and we confirm the hard-easy effect assumption. A new result emerges from our analysis: we show that confidence has a good and bad component meaning that it does not always predict success. A moderate self-confidence has a positive impact on success whereas a level of confidence that is too high can lead to failure. We also confirm the aspiration upgrade after a positive update of abilities.

4. For example, the initial aspiration for the GT subjects is set to 0. Hence, if they succeed level 1 they are already above their initial aspiration. They may then upgrade their aspiration by taking into account the fact that they fulfilled their initial goal and estimate they have enough ability to continue and try to succeed level 2,

This fact can be seen on our gain and intermediate treatment subjects: for IT (GT) high (low) ability, reaching level 2 (level 1) (which represent their induced aspiration level) modifies their aspiration by motivating them to continue and succeed further. The opposite effect is found for low ability IT subject.

The remaining of the paper is organized as follows. Section 2 will describe the experimental design. Section 3 and 4 respectively present the theoretical notions of aspiration and confidence. Section 5 focuses on the decision model. Section 6 estimates the model showing a confidence analysis and the aspiration upgrade analysis. Finally section 7 discusses the results and section 8 concludes.

2. The experiment

2.1. *The design*

The design is inspired by Page, Levy-Garboua and Montmarquette (2007) and aims at reproducing a choice process with three levels of difficulty. The decision to continue to a higher difficulty level only arises if a level is succeeded. When failing a level, the subject has to quit the experiment⁵. We emphasize the selective nature of the process rather than its learning function.

Participants perform a real-effort task for which they get paid according to their degree of success. The task consists in solving anagrams ranked in three levels of increasing difficulty. It is performed during a maximum of 15 rounds lasting no more than 8 minutes each. These 15 rounds are structured in three successive levels of increasing difficulty, designated respectively as level 1, 2 and 3. Participants are successful at

5. During an experimental session, subjects quit the lab at different timing according to the moment they fail, succeed, or decide to stop the experiment. It could be argued that this could create contamination behaviors. However, the task was completely individual, there were no interaction between subjects, and they could not know if a subject was leaving the room because of failure or because she decided to stop. Plus, concentration was such that at the end of the sessions many subjects did not even realize that some had already left the room.

one level when they manage to decode at least $2/3$ of the anagrams displayed. Level 1 consists of 9 rounds of low difficulty (6 anagrams per round to be solved in no more than eight minutes), which leaves ample time to learn the task. The next levels comprise 3 rounds each with a greater number of anagrams of same difficulty to be solved under the same time limit. Figure 1.7 in the appendix shows the exact screen faced by subjects during step 1 of level 1 (six anagrams per stage). The anagrams were created from a list of 2000 six-letter words that are most used in French. We then randomly chose 114 words, for which we randomly mixed the six letters. All subjects must go through the first level, but they have an option to leave the game or to continue to the next higher level after succeeding level 1 or level 2. Subjects who fail to meet one level's requirement must drop-out of the game. The calibration of anagrams has been made such that we obtain a high enough success rate among low ability subjects, and almost a 100% success among high ability subjects. This goal is completely fulfilled as 69% of low-ability and 95% of high ability subjects succeeded level 1 (the ability variable will be carefully described in section 6).

The payoffs were framed according to gain, loss or both to induce an initial aspiration. We designed three treatments, one for each framing. Table 1.1 presents the different payoffs among treatments, according to the success or the failure at each level. We thus simulate three aspiration levels with three levels of endowments that modulate the perception of the payoff (loss or gain) and thus the reference point (see the prospect theory of [Kahneman and Tversky \(1979\)](#)).

The decision that a subject has to make at the end of each completed level becomes a choice between a sure payoff (if she decides to stop and to leave) and a lottery with a probability p of winning the next level and a probability $1 - p$ of failing the level (if she decides to continue). In this manner, the whole experiment can be summarized as a decision process (see the appendix figure 1.8 to have a better idea of the decision tree

		Treatments		
Levels		Gain	Intermediate	Loss
	Initial endowment	0€	20€	35€
Level 1	Fail	2€	-18€	-33€
	Pass and stop	10€	-10€	-25€
Level 2	Fail	4€	-16€	-31€
	Pass and stop	20€	-0€	-15€
Level 3	Fail	11€	-9€	-24€
	Pass and stop	35€	+15€	-0€

Table 1.1: Initial endowments and additional gains and losses

which subjects faced). In fact, if a subject is not sure of her ability, continuing is risky as if it leads to failure at the level, then the payoff is smaller than if the subject had chosen to stop. This represents the opportunity cost of continuing.

Moreover, at the end of the first level, and only if they succeed and decide to continue, subjects have to choose between two curricula; the "Wall" and the "Hill"⁶. Payoffs for both curricula are the same, only the difficulty with the number of anagrams to be solved at each level changes. Subjects have to choose between the curricula once they completed level 1 and decide whether to continue further. In the "Wall" condition, the difficulty increases sharply at level 2 (with 10 anagrams per stage so that subjects consequently have to solve 20 anagrams to pass the level), but remains constant at level 3 with the same conditions. In the "Hill" condition, the difficulty always rises from one level to the next, marginally at level 2, then sharply at level 3. Instead of having 6 anagrams per stage, subjects have to solve 8 anagrams and thus 16 anagrams to clear the second level. If they decide to take the third level, the difficulty increases a lot: there are now 12 anagrams per stage, and 24 anagrams have to be solved to succeed the last level. By the end of the experiment, the required number of anagrams is the same for the "Wall" and "Hill" conditions. Table 1.2 reports the number of anagrams

6. This feature tries to represent different types of tracks one can take during a career or educational path for example. In these cases, there is never only one possible track.

that have to be solved, by level and curriculum.

Levels	Level 1	Level 2		Level 3	
Chosen curriculum		Hill	Wall	Hill	Wall
Number of stages	9	3	3	3	3
Number of anagrams by stage	6	8	10	12	10
Minimum number of anagrams to succeed	36	16	20	24	20

Table 1.2: Number of anagrams that have to be solved by level and curriculum

We assume that by observing the path subjects choose, this allows us to check for the subjects' expected attained level and also to see if they make an accurate estimation of their abilities. This choice should maximize the expected success; a subject unable to reach level 2 by the "Hill" track should stop at level 1; and a subject who feels she is able to solve between 6 and 8 anagrams per round but no more should choose the "Hill" track. Of course, if she passes level 2 successfully, she can update expected payoff by continuing to level 3 but realizing that difficulty still increases. On the other hand, the "Wall" track was designed for subjects who are confident enough in succeeding level 3, and who seek to get the payoff associated to this level. Indeed, as the number of anagrams required to pass level 2 during the "Wall" track is the same at level 3, the probability of succeeding level 3, conditional on succeeding level 2 tends to 1. To avoid any obvious parallels between the "Hill" and "Wall", we named both tracks A and B and reversed their order of presentation in half of the sessions, to control for order effects.

2.2. Confidence elicitation

As our model is based on confidence theories, we elicit confidence to test if participants are subject to both estimation and miscalibration biases. Confidence is elicited twice: once just after step 4 of level 1 (so during level 1) and once at the end of level 1 only for subjects who succeed level 1 and decide to continue to level 2 (thus just before

level 2). Each time, subjects are asked to declare a percentage probability (between 0 and 100) of succeeding each level. We denote these confidence level as $q^{\ell,4}$ and $q^{\ell,9}$, ℓ being the considered level (1, 2 or 3 for $q^{\ell,4}$, and 2 or 3 for $q^{\ell,9}$), and 4 or 9 the stage when confidence is elicited. Overall, we get either 3 confidence levels for subjects who fail or stop at level 1, and 5 confidence levels for those who continue to level 2. Figure 1.1 illustrates the different steps of the experiment⁷.

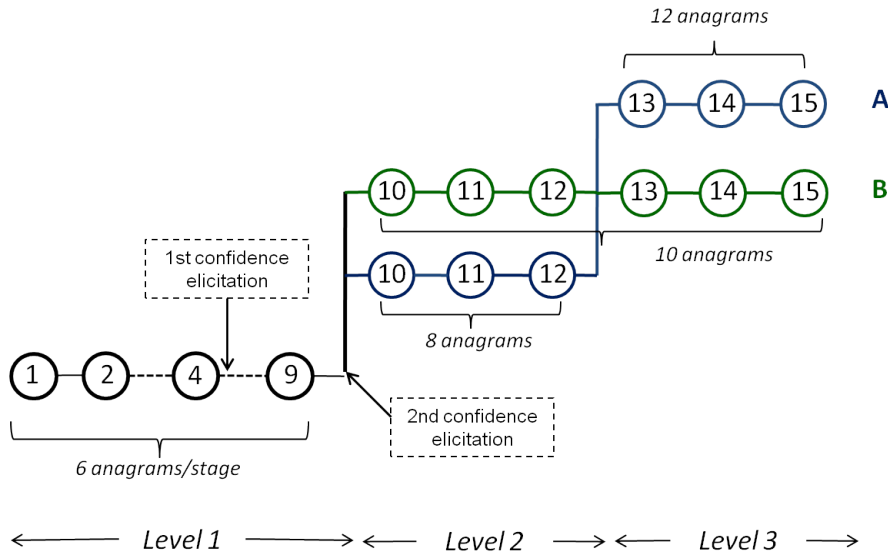


Figure 1.1: Experiment scheme

Confidence elicitation is not incentivized and the instructions (that can be found at the end of the appendix) do not tell about confidence elicitation. Subjects do not know in advance that they will have to answer this question. First, this removes the possibility for subjects to think carefully about their confidence level too much in advance. Second, subjects are asked to report their true beliefs sincerely. The true beliefs dictate the actual observed behavior (continuing, succeeding) and this is incentivized by the different possible payoffs. Incentivizing beliefs on two successive occasions can induce risk-averse subjects to diversify their reported estimates as a hedge against the

⁷. A corresponds to the "Hill" track, B to the "Wall" track.

risk of prediction error (Armantier and Treich, 2013). Moreover, self-reported methods have been widely used and validated by psychologist and neuroscientists. Careful and recent comparisons of self-reports methods with quadratic scoring rule found that it performed as well (Clark and Friesen, 2009) or even better (Hollard, Massoni and Vergnaud, 2010) than the quadratic scoring rule. Last, the experiment could last up to 2 hours and 30 minutes (for a subject who reaches level 3)⁸ and is quite complex, so incentivizing the belief with a specific rule would have made the design even more complex. This could create a focusing on confidence itself rather than on the whole decision process. Using self-reported measures seems to be here more appropriate as it is faster and simpler. The results on the analyse of confidence and success presented in later sections support this choice.

2.3. Descriptive statistics

We ran a total of 14 sessions. Half of them were conducted in Montreal (*Cirano*) and the other half in Paris (at the *Laboratoire d'Economie Expérimentale de Paris*). One session corresponded to one treatment (see table 1.11 in appendix). We always read the instructions aloud before starting the experimental program. Subjects were free to ask as many questions as they wanted. To make sure they perfectly understood the experiment they had to answer a comprehension questionnaire (11 questions) before starting to solve the anagrams⁹. They could not go on until every answer to these questions was correct. After this questionnaire, subjects were asked some personal questions about their age, their gender, their level of education, their mother tongue, their frequency of playing Scrabble or doing crosswords etc... These questions were asked to make sure that if we found any significant differences between our groups, we would be able to control for them in our estimations. After having answered all

8. The duration of the experiment was quite long, but when subjects subscribed to the experiment they already knew that they had to be available for 2h30. This avoids the problem that subjects quit the experiment because they have something else planned rather than because of the experiment itself.

9. The comprehension questionnaire can be found in the appendix after the instructions.

of these questions, subjects could start level 1 of the experiment. Decisions always remained private, and communication between subjects was not allowed. When they finished their experimental session, and before being paid off, subjects had to answer to two subjective questions about their level of stress and their satisfaction on a 7 points Likert-scale (see table 1.3 for descriptive statistics).

There were a total of 243 participants and earned on average 23.3€, including a participation fee of €7. In Montreal, all payments were made in Canadian dollars, based on purchasing power parity. However, we only use euros in the rest of the paper for the sake of simplicity.

Variables	Modality	GT	IT	LT	difference
Gender	Men	48,7%	47,6%	55,6%	ns
Age		26,5	27,2	26,6	ns
Level attained		1,5	1,7	1,6	ns
Mother tongue ¹⁰	French	75,0%	73,2%	56,8%	** <i>IT≠LT</i> ** <i>GT≠LT</i>
Prior participation in an experiment	Yes	81,2%	84,1%	77,8%	ns
Educational level ¹¹	Bac +3 to bac+5 or more	61,2%	63,4%	70,4%	ns
Occupation	Work or study	90,0%	95,1%	95,1%	ns
Risk Aversion	Yes	66,3%	80,5%	74,1%	** <i>GT≠IT</i>
Cross words	Occasionally or regularly	26,3%	34,2%	24,7%	ns
Scrabble	Occasionally or regularly	12,5%	28,1%	16,1%	* <i>IT≠LT</i> *** <i>IT≠GT</i>
Number of comprehension mistakes		2,1	2,6	2,2	ns
Stress ¹²	scale from 1 to 7	4,9	4,9	4,8	ns
Satisfaction	scale from 1 to 7	3,8	4,4	4	* <i>GT≠IT</i>
N		80	82	81	ns

Table 1.3: Descriptive statistics
Significance levels of two-tailed z-test or t-test: ** 5% *** 1%

Even though French is the first language in Montreal, some English speakers live in this city and are less familiar with French word games as anagrams, scrabble or cross words. 81.9% of the Parisian participants speak French as their mother tongue against

10. We started with 12 sessions but we had to add two additional sessions in Paris, as we had significantly less subjects in the GT and IT. This explains why we find more French speakers in the GT and IT.

11. Equivalent Canadian educational levels are "Etudes secondaires to diplômes d'études professionnelles" and "Etudes collégiales to études universitaires".

12. A scale between 1 and 7. The smaller the number, the less stressed or satisfied the subject.

51.8% among the participants from Montreal. We can expect to find differences among these two types of population in terms of control variables and abilities. However, no significant differences were found between these two groups for the different control variables presented in table 1.3. In terms of success, there is no significantly more Parisians than Montrealers who succeed better at the three levels¹³. The fact that we have two different populations can be seen as an advantage as it implies a greater robustness of the results.

Concerning risk aversion, that has an important role to play in our model, we ask subjects a non-incentivized question. They had to chose between a sure payment of €5 and a €10 payoff with an uncertain probability. This question is in our mind sufficient to have an idea of the proportion of risk averse subjects in our sample. The primary aim of the paper is not to measure risk aversion *per se*, but rather to analyze its role during the decision process. Table 1.3 shows that most of our subjects are risk averse.

2.4. *Conditional success rate and choice of curriculum*

Before going into deeper details of the model and estimations, we present in this subsection success and failure rates conditional on the decision to continue to the next difficulty level (see table 1.4).

Gain treatment subjects have a higher rate of failure at level 1 and 3 (although not significant). Subjects from the intermediate treatment stop more in proportion when they succeed level 1 but stop less when having succeeded level 2. They are also more represented in level 3 success (even though differences are not significant) as they significantly succeed better than the two other treatments' subjects.

Concerning the choice of curriculum, we expect that participants with lower abilities

13. At level 1 success rates yields 82.2% for Parisians vs. 81.2% for Montrealers, p-value from a two-sided z-test=0.75. Total level 2 success: 53.4% vs. 44.5%, p-value=0.17. Total level 3 success: 28.6% vs. 32.7%, p-value=0.48.

14. The difference is also significant between GT and LT at a 10% level.

	GT		IT		LT
	100		100		100
Fail level 1	21.3		15.9		17.3
Succeed level 1 and stop	7.5	↔ *	14.6		9.8
Continue and fail level 2	22.5		18.3		24.7
Succeed level 2 and stop	13.8		8.5	↔ **	18.5
Continue and fail level 3	8.8	↔ * ¹⁴	3.6		3.7
Succeed level 3	26.4		39.1		26

Significance levels of two-tailed z-test: * 10% ** 5% *** 1%

Table 1.4: Success and failure rates, by treatment (in %)

should choose "Hill". In fact, as low ability participant they should not expect to succeed level 3. By choosing "Hill" they would be more likely to succeed at least the second level than if they had chosen "Wall", which provides them the opportunity to earn more with less effort. Moreover, as we assume an effect of the treatment on this same choice, we would expect that subjects having a lower level of aspirations, that is in the GT and IT, would have a higher probability to choose "Hill". Indeed, their achievement goal being lower, they might seek to reach level 2. Figure 1.2 reports the curriculum choice proportions according to the treatments. Contrary to what we expected, overall, subjects tend to choose significantly more the "Wall" path rather than the "Hill" path and the rates are not significantly different according to the treatments. As the choice of "Hill" and "Wall" does not differ among treatments and seems not to be truly taken into account by our subjects, we exclude this feature from our analysis and will only consider it as a control variable¹⁵.

15. Askari, Gazel and Lévy-Garboua (2013) present a similar experiment which is more focused on this paths choice.

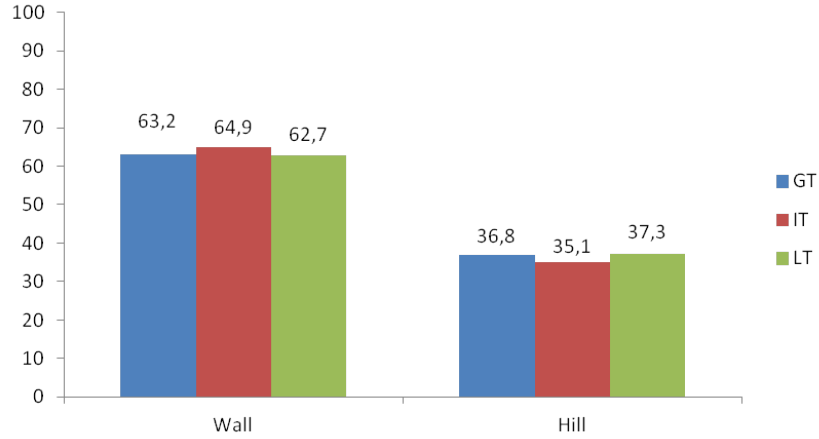


Figure 1.2: Proportion of choices between both track according to treatments (in %).

3. The aspiration level

We develop a behavioral model of the choice of effort at the successive levels of increasing difficulty and of the decision to continue the task at a higher level. This decision relies on the subjective probability of success, which depends on the level of effort the agent will provide, the ability level, the true probability of success (fully determined by ability and effort) and the aspiration level. We consider two types of agents of high and low ability $i = \{H, L\}$, and three levels of difficulty $\ell = \{1, 2, 3\}$.

\tilde{k} is the true ability level and is a continuous variable¹⁶. We make the hypothesis that individuals do not know their true level of abilities *ex ante* and make their decisions based on their perceived abilities. According to the available information they already have (past successes, training in the task or in related tasks) they perceive their abilities $E(k)$ such that $\tilde{k} \neq E(k)$. Let then e be the effort provided by the agent. Effort is unobservable but effort variations can be observed through confidence and success. Note that the decision to continue is only considered for agents who believe they can provide an effort that will allow them to potentially succeed. If their perceived abili-

¹⁶. Among each type, we allow for additional heterogeneity in abilities

ties and effort needed is below this threshold, they will decide to stop the game anyway.

As explained in the introduction, aspiration, denoted as A in the model, is considered here as a desired goal, equal or greater than the status quo, that agents adopt as their reference at least equal or greater than the initial level. Each agent has a reference-dependent utility function $u(y_\ell - A)$, $u' > 0$ (Kahneman and Tversky, 1979, Köszegi and Rabin, 2007).

Aspirations are partly endogenous and represent ambition: once initial aspirations are fulfilled (mainly because of enough abilities), people are inclined to upgrade them to a higher level if the latter exists (Hoppe, 1976), even though it may be more difficult to reach. We consider then that an individual who has ambitions is someone having aspiration above the status quo level.

Aspirations are expressed by the level A_ℓ that an agent wishes to pass when she stands at level $\ell - 1$. We postulate that aspirations, that is, the new goal to be reached, are governed by the agent's estimated ability level $E_\ell k$ and by her ambition at this stage. Both estimated ability level and ambitions are endogenous. The level of aspiration reached corresponds to a level of payoff introduced in the utility function of the agent. An exogenous part corresponds to the initial aspirations A_0 induced by the endowment in our experiment.

$$A_0 = \begin{cases} 0 & \text{For the Gain Treatment} \\ 2 & \text{For the Intermediate Treatment} \\ 3 & \text{For the Loss Treatment} \end{cases}$$

We posit that:

$$A_1 = \text{Max}(f(E_1k), A_0) \text{ if } f(E_1k) = \begin{cases} 0 & \text{if } 0 \leq E_1k < k_2 \\ 2 & \text{if } k_2 \leq E_1k < k_3 \\ 3 & \text{if } k_3 \leq E_1k < K \end{cases}$$

$$A_2 = \text{Max}(f(E_2k), A_1)$$

$$A_3 = \text{Max}(f(E_3k), A_2)$$

The aspiration level is either stable or increasing as long as agents do not drop out:

$$A_3 \geq A_2 \geq A_1.$$

$f(k)$ is a non-negative increasing step function (represented in figure 1.3) such that $f : [k_2; K] \rightarrow [2, 3]$, k_ℓ being the minimum ability level required to succeed level ℓ , such that:

$$\begin{cases} f(k) = 2, & \text{if } k_2 \leq k < k_3 \\ f(k) = 3, & \text{if } k_3 \leq k \leq K \end{cases}$$

Consider our three experimental treatment groups to illustrate the aspirations and ability update process:

- For the Gain Treatment ($A_0 = 0$), we always have:

$$A_\ell = f(E_\ell k) \text{ for all } \ell = \{1, 2, 3\}$$

For low-ability subject in this treatment, a specific case arises. If they manage to pass level 1, then E_1k is updated to $E_2k > E_1k$. Indeed, when starting the game, they have an estimation of their ability for level 1 (E_1k). When succeeding level 1 they increase their ability estimation for level 2 (E_2k). This implies: $A_2 = \text{max}(2, A_0) = 2$, if $E_2k \geq k_2$.

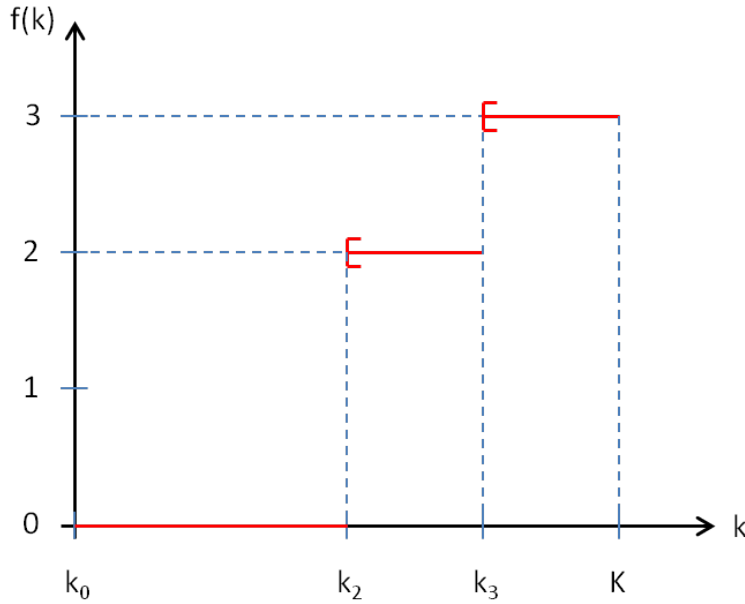


Figure 1.3

- For the Intermediate Treatment ($A_0 = 2$):

1) If $E_1k \geq k_3$ (meaning that they start the game with a high estimation of their ability), then $f(E_1k) = 3$ and $A_0 = 2$, hence, $A_1 = 3$.

If they pass level 1, $f(E_2k) = f(E_1k) = 3$.

If they pass level 2, $f(E_3k) = f(E_2k) = 3$.

2) If $k_2 \leq E_1k < k_3$, $f(E_1k) = 2 = A_0 \Rightarrow A_1 = 2$

3) If $0 \leq E_1k < k_2$, then $A_1 = A_0 = 2$.

Once they succeed level 2, then $E_3k > E_2k \geq E_1k$. If $E_3k \geq k_3$, they continue to level 3, hence $f(E_3k) = 3 \Rightarrow A_3 = 3$. In this case, IT subjects have high abilities but are under-confident. They realize their ability level as they succeed level 2 and thus update them, implying an aspiration update to level 3, greater than the induced aspiration level.

- For the Loss Treatment:

$A_0 = 3$. Hence, $A_\ell = \max(f(E_\ell k), A_\ell) \geq A_\ell, \forall \ell$.

$$\Rightarrow A_3 = A_2 = A_1 = 3$$

Subjects in the loss treatment always keep a high level of aspiration irrespective of their true ability and confidence.

Proposition 1 *Agents upgrade their aspiration level with positive probability if they manage to reach their prior aspiration level.*

Aspiration will then play a large role when interacted with the ability. This effect mainly affects people who can revise their aspiration levels along the different difficulty levels, that is the low-ability GT and high-ability IT subject. This is largely discussed in section 6.2. Moreover, the proportion of agents who upgrade their ambition will depend on the induced aspiration level.

4. Self-confidence

We set the true probability of success as being $\bar{q}_\ell(e, \tilde{k})$ where $e \in [0, e_{max}]$ and with $\bar{q}'_k \ell(e, \tilde{k}) > 0$, $\bar{q}'_e \ell(e, \tilde{k}) > 0$ and, $\bar{q}''_{ee}(e, \tilde{k}) < 0$, $\bar{q}''_{kk}(e, \tilde{k}) < 0$. The sign of $\bar{q}''_{e\tilde{k}}(e, \tilde{k})$ depends on the complementarity (>0) or substitutability (<0) between effort and perceived ability.

The level of effort the agent will provide and her true ability will determine the true probability of success. Moreover, the quantity of effort provided to succeed will vary with time and with the difficulty of level ℓ .

Confidence q is defined as a subjective probability of success (ex ante). That self-confidence covaries with the true probability of success is a minimum consistency requirement, and the probability of success is obviously conditioned by the choice of effort. For example, if an agent decides not to provide any effort, confidence in succeeding the task may be driven to zero¹⁷. Globally, the beliefs in succeeding the task

17. Another case is also possible: a highly able individual may consider that increasing the level of effort will not increase the probability of success as she thinks she already has enough ability to succeed, which will not impact confidence. The opposite reasoning can be made for low ability people.

will vary with the effort level provided. Confidence may differ from the true probability of success for two reasons: (i) because agents have an imperfect knowledge of their own abilities \tilde{k} (estimation bias) and (ii) because agents suffer from a systematic calibration bias when they make a choice (even if they know their abilities).

4.1. Estimation bias

If individuals perfectly knew their ability level, their confidence would be equal to the true probability of success conditional on their abilities. Individuals experience an *estimation bias* on their ability level and thus on their probability of success. This bias is $E(\bar{q}) - \bar{q}$, with $\bar{q} = \bar{q}(e, \tilde{k})$ and $E(\bar{q}) = \bar{q}(e, E(k)) \in [0; 1]$ ¹⁸.

Agents are assumed to update their abilities with experience in a Bayesian fashion. The estimated ability will rise after a better-than-expected performance and decline after a worse-than-expected performance. Good and bad performances will be equally recorded and weighted. Thus, the estimation bias goes down to zero in the long run.

4.2. Miscalibration bias

In the short run, however, Bayesian people may miscalibrate their own probability of success through a miscalibration parameter $1 - \mu_i \in [0; 1]$ even if they have an unbiased estimate of their own ability. The miscalibration bias is then $q_i - E(\bar{q})$ for $i = \{H, L\}$. This can occur if they are uncertain of the true probability of success because they can be misled by emotions triggered by their doubt. However, the direction of doubt is entirely different depending on whether their prior estimate led them to believe that they would fail or that they would succeed. We distinguish *miscalibration* among both types of agents:

18. Under either assumption, we assert that ability increases the value of effort. Measuring ability in efficiency units, the output function of ability and effort is either additive ($k + e$) as ability can add value to the effort, or multiplicative ($k.e$). Under any of these assumptions, $E(\bar{q}) = \bar{q}(e, E(k))$.

- For low ability individuals:

$$q_L = \mu_L E(\bar{q}) + (1 - \mu_L).1 \quad (1.1)$$

This equation implies that the calibration bias is:

$$q_L - E(\bar{q}) = (1 - \mu_L)(1 - E(\bar{q})) > 0 \text{ if } 1 - \mu_L > 0 \quad (1.2)$$

Low-ability subjects are hence overcalibrated. They should normally fail at middle or higher levels. However, their perception of a possible success leads to overweighting their subjective probability of success, i.e. overconfidence. Thus, even though low-ability agents should give up a task, they are overconfident and are thus tempted by the returns to success (cf. equation 1.1). One can notice that if one expects no success, that is $E(\bar{q}) = 0$, then confidence $q_L = 1 - \mu_L > 0$. This means that low-ability individuals always exhibit a positive bottom confidence which is in line with the Dunning-Kruger effect (they overestimate their abilities.)

- For high ability individuals:

$$q_H = \mu_H E(\bar{q}) + (1 - \mu_H).0 \quad (1.3)$$

This equation implies that the calibration bias is:

$$q_H - E(\bar{q}) = -(1 - \mu_H)E(\bar{q}) < 0 \text{ if } \mu_H < 1 \quad (1.4)$$

High-ability subjects are undercalibrated. They should normally succeed at middle or higher levels. However, the perception of a possible failure leads to underweighting their subjective probability of success, i.e. underconfidence (cf. equation 1.3). If high ability agents expect a 1 probability of success (i.e. $E(\bar{q}) = 1$),

their confidence $q = \mu_H \leq E(\bar{q})$. Hence, even though they expect a success through an accurate estimation of their abilities, they always exhibit a ceiling confidence lower than 1 confirming as well the Dunning-Kruger effect where high ability subjects underestimate their abilities.

4.3. *Synthesis on confidence*

In order to generalize equations 1.1 and 1.3, we introduce a dummy variable $D(L)$ such that it is equal to one for low ability agents, and to zero for high ability agents. The miscalibration parameter $(1 - \mu_i)$ and the estimation bias $(E(\bar{q}) - \bar{q})$ can be grouped into this unique equation of confidence:

$$q_i = \mu_i E(\bar{q}) + (1 - \mu_i) D(L) = \bar{q} + \mu_i (E(\bar{q}) - \bar{q}) + (1 - \mu_i) (D(L) - \bar{q}) \quad (1.5)$$

Equations 1.1, 1.3 and 1.5 are a formalisation of the Dunning-Kruger effect. We can see that if both of the biases disappear, that is $E(\bar{q}) - \bar{q} = 0$ and $1 - \mu_i = 0$, then confidence is equal to true success.

4.3.1. **Total confidence bias**

The total confidence bias is thus given by $q_i - \bar{q}$, that is the difference between confidence and the true probability of success. If it is positive (negative) then the agent is overconfident (underconfident), if it is equal to zero the agent is accurate. It can be written as:

$$q_i - \bar{q} = (E(\bar{q}) - \bar{q}) + (q_i - E(\bar{q})) \quad (1.6)$$

This yields the following proposition:

Proposition 2 *The confidence bias can be decomposed additively in the ability estimation bias $(E(\bar{q}) - \bar{q})$ and the calibration bias $(q_i - E(\bar{q}))$.*

4.3.2. Confidence bias for low-ability agents

In this section we analyze the confidence bias of low-ability agents, implying $D(L) =$

1. By using equation 1.5, the confidence bias is:

$$q_L - \bar{q} = \mu_L(E(\bar{q}) - \bar{q}) + (1 - \mu_L)(1 - \bar{q})$$

Note that $1 - \bar{q} \gg E(\bar{q}) - \bar{q}$. Hence, $q_L - \bar{q} > E(\bar{q}) - \bar{q}$ if $1 - \mu_L > 0$.

The total bias can be rewritten as: $q_L - \bar{q} = 1 - \bar{q} - \mu_L(1 - E(\bar{q}))$.

Low ability agents exhibit over-confidence, that is

$$\begin{aligned} q_L > \bar{q}_L &\text{ iff } 1 - \bar{q} > \mu_L(1 - E(\bar{q})) \\ \text{or, } q_L > \bar{q}_L &\text{ iff } E(\bar{q}) - \bar{q} > -\frac{1-\mu_L}{\mu_L}(1 - \bar{q}) \end{aligned}$$

Proposition 3 *Low-ability subjects are over-calibrated, i.e. $q_L > E(\bar{q})$. The calibration bias is proportional to the miscalibration factor $(1 - \mu_L)$ and gets larger as ability gets lower (cf. equation 1.2).*

Low-ability subjects are over-confident iff ability-estimation is not too biased downward, i.e. $E(\bar{q}) - \bar{q} > -\frac{1-\mu_L}{\mu_L}(1 - \bar{q})$.

4.3.3. Confidence bias for high-ability agents

For high-ability agents, $D(L) = 0$. By using equation 1.5, the confidence bias is:

$$q_H - \bar{q} = \mu_H(E(\bar{q}) - \bar{q}) - (1 - \mu_H)\bar{q}$$

The total bias can be rewritten as (cf. equation 1.3): $q_H - \bar{q} = \mu_H E(\bar{q}) - \bar{q} < E(\bar{q})$ (with $\mu_H < 1$).

High ability agents exhibit under-confidence, that is:

$$\begin{aligned} q_H < \bar{q} &\text{ iff } \mu_H E(\bar{q}) < \bar{q} \\ \text{or, } q_H < \bar{q} &\text{ iff } E(\bar{q}) - \bar{q} < \frac{1-\mu_H}{\mu_H}\bar{q} \end{aligned}$$

Proposition 4 *High-ability subjects are under-calibrated, i.e. $q_H < E(\bar{q})$. The calibration bias is proportional to the miscalibration factor $(1 - \mu_H)$ and gets larger as ability gets higher (in absolute value, cf. equation 1.4).*

High-ability subjects are under-confident iff ability-estimation is not too biased upward, i.e. $E(\bar{q}) - \bar{q} < \frac{1-\mu_H}{\mu_H}\bar{q}$.

Corollary 1 *If ability estimation is unbiased (i.e. $E(\bar{q}) = \bar{q}$), low-ability subjects are always overconfident, and high-ability subjects are always under-confident.*

4.4. Aspiration and the miscalibration parameter

An implication would be that considering the low ability agents' confidence level, the perception of a possible success is salient when the aspiration level is high because the signal of success is consonant with the subject's ambitious goal. It would become less salient if the aspiration level is low because the signal of success is then dissonant with the subject's lack of ambition. Consequently, this emotional signal will be weighted more heavily in equation 1.1 and the overconfidence bias will increase.

Considering the high ability agents' confidence level, the perception of a possible failure is never consonant with the aspiration level, even when ambition is lacking. This perception is thus likely to be less salient than the perception of success for low-ability individuals having high aspirations; and it should not vary much with the aspiration level. Consequently, this emotional signal will not be weighted very heavily in equation 1.3 and the underconfidence bias will vary little with aspirations. This reasoning yield an assumption on the relationship between aspirations and the miscalibration bias:

Assumption 1 *The miscalibration parameter $1 - \mu_L$ increases with the aspiration level for low-ability subjects, $\frac{\partial 1 - \mu_L}{\partial A} > 0$.*

The miscalibration parameter $1 - \mu_H$ varies little with the aspiration level for high-ability subjects, $\frac{\partial 1 - \mu_H}{\partial A}$ is almost equal to 0

5. The decision model

Based on our experimental design with three levels of increasing difficulty, when succeeding the first or second level, the agent can decide whether to engage in the risk-taking activity, that is, to stop or continue to the next level. If she stops at level $\ell - 1$, for $\ell = \{1, 2\}$ she gets a sure outcome $y_{\ell-1}$; and if she continues, she gets a payoff: y_ℓ if the level ℓ is succeeded, $y_{\ell,0}$ if the level is failed such that $y_\ell > y_{\ell-1} > y_{\ell,0}$. Once agents have estimated their ability, our behavioral model relates effort, self-confidence, risk-taking and success to the individual level of aspirations and propensity to miscalibrate. Recalling that the success rate, calibration parameter, ability estimation and self-confidence are given by:

- Success rate: $\bar{q}_\ell \equiv \bar{q}_\ell(e, k)$
- Calibration rate: $\mu_i \equiv \mu_i(A)$.
- Ability estimation: $E(\bar{q}) = \bar{q}(e, E(k))$.
- Confidence: $q_i \equiv \mu_i E(\bar{q}) + (1 - \mu_i)D(L)$

We consider a cost of effort $c(e)$ such that $c'(e) > 0$ and $c''(e) > 0$.

The agent maximises her expected utility:

$$\max_{e \in [0, e_{max}]} V_\ell(e) = q_i(e, E(k))u(y_\ell - A) + (1 - q_i(e, E(k)))u(y_{\ell,0} - A) - c(e) \quad (1.7)$$

The first order condition for an interior solution yields:

$$\begin{aligned} \frac{\partial V_\ell(e)}{\partial e} &= q'(e, E(k))[u(y_\ell - A) - u(y_{\ell,0} - A)] - c'(e) = 0 & (1.8) \\ &\iff q'_e(e, E(k))W_\ell(A) = c'(e) \\ &\iff \mu_i(A)\bar{q}'_e(e, E(k))W_\ell(A) = c'(e) \end{aligned}$$

With $W_\ell(A) = u(y_\ell - A) - u(y_{\ell,0} - A) > 0$

The agent decides then to continue to the higher level $\ell = \{2, 3\}$ iff her expected utility $V_\ell(e)$ is greater than the utility obtained at the previous level $\ell - 1$ that is:

$$V_\ell(e) = q_i(e, E(k))u(y_\ell - A) + (1 - q_i(e, E(k)))u(y_{\ell,0} - A) - c(e) > u(y_{\ell-1} - A) \quad (1.9)$$

Optimal level of effort is determined so that expected utility can be maximized and compared to the previous level utility, allowing to make the decision to continue or not. Figure 1.4 summarizes the different steps of decision and updating processes in this dynamic context. After each success, the updating process arises and yields a new optimal level of effort determining again the future decision to continue to the next level.

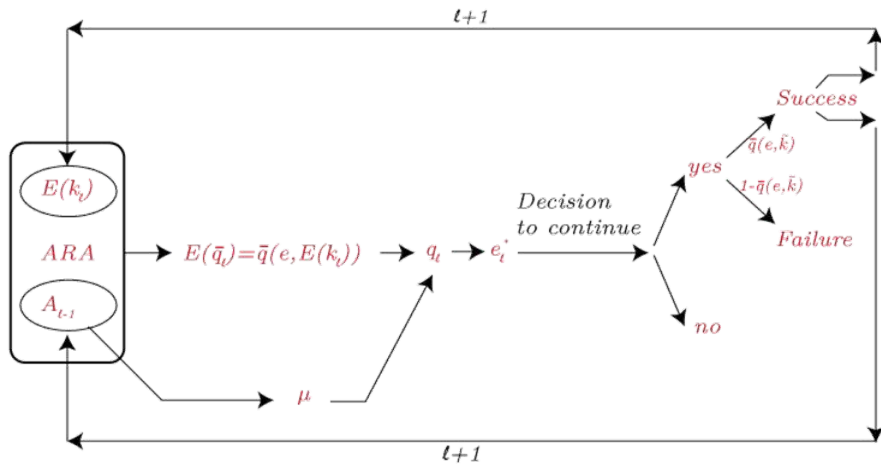


Figure 1.4: Summary of the decision and updating process

In the following subsections we analyse the variation of effort according to $E(k)$, aspirations A , risk aversion, and to the calibration parameter μ . Derivatives of equation 1.8 with respect to these variables can be found in the appendix (as proofs).

5.1. Effort and estimation bias

The sign of $\frac{\partial e^*}{\partial E(k)}$, that is the effect of perceived abilities on effort will depend on the performance technology. If effort and abilities are complements in the production of success, then the effect of abilities on effort is positive. On the contrary, if they are substitutes, higher abilities would lead to a lower effort.

The most plausible interpretation is the first one. If one has high abilities, one is willing to succeed and thus will exert a higher level of effort, especially when a higher level has to be attained that is a higher expected payoff. However, the other case can happen if, for example, one's abilities are just enough to pass a level, and no increase in effort is needed to improve performance.

Proposition 5 *A positive ability-estimation bias enhances effort, self-confidence, and performance if ability and effort are complementary factors of success. The reverse conclusion holds if ability and effort are substitutes.*

5.2. Effort and miscalibration

$\frac{\partial e^*}{\partial \mu(A)}$ is always positive. So, the higher the μ , meaning the closer to the true probability of success is your confidence, the higher the level of effort exerted. In other words, a higher miscalibration parameter $(1 - \mu)$ implies a smaller effort, confidence and performance.

Proposition 6 *The miscalibration bias increases self-confidence but always reduces effort and success.*

5.3. Effort and aspiration levels

The reference-dependent utility function includes the level of aspirations. Hence, the sign of $\frac{\partial e^*}{\partial \mu(A)}$, that is the impact of aspiration levels on the effort provided, will thus

depend on risk aversion. Consequently, the sign of $\frac{\partial e^*}{\partial A}$ depends on the signs of W'_i and $\frac{\partial \mu_i(A)}{\partial A}$ (cf. assumption 1).

Table 3 displays the effect of the aspiration level on effort by considering risk preferences of the agent.

	Risk Averse ($\frac{\partial W_i(A)}{\partial A} > 0$)	Risk Lover ($\frac{\partial W_i(A)}{\partial A} < 0$)
If $i = L$, $\frac{\partial \mu_L}{\partial A} < 0$	sign of $\frac{\partial e^*}{\partial A}$ indeterminate	$\frac{\partial e^*}{\partial A} < 0$
If $i = H$, $\frac{\partial \mu_H}{\partial A} = 0$	$\frac{\partial e^*}{\partial A} > 0$	$\frac{\partial e^*}{\partial A} < 0$

Table 1.5: Effect of the aspirations on effort depending on risk preferences.

Proposition 7 *If a majority of subjects are risk-averse (confirmed in table 2.7 for our sample), the aspiration level should have little effect on effort and, therefore, on confidence for low-ability subjects. In contrast, the aspiration level should have a substantial positive effect on effort and, therefore, on confidence for high-ability subjects.*

The smaller the proportion of risk-averse subjects in the population, the less positive or more negative is the effect of aspirations on effort and confidence.

6. Estimation of the model

Our experimental data allows us to test our model. The next subsections describe the estimations of the miscalibration parameter and show how aspirations, abilities and confidence explain the level of effort exerted (unobserved but the true probability of success gives information on the variation of effort).

In order to distinguish both ability levels on our experimental subjects, we create a performance variable called *the average time cost*, indicating the average time a subject takes to solve one anagram during the first four stages. This variable is quite exogenous

in the way that it can be considered as the cognitive ability our subjects already possess. We do not consider any learning effect that may exist during the experiment. High ability subjects are situated below the median of this variable, and low ability subjects are above the median. We will continue to refer to high and low ability subjects based on this variable, and this specification will always remain the same. The proportions of high and low performers in the various treatments are not significantly different¹⁹.

6.1. Estimation of the miscalibration parameter $1 - \mu$

In this subsection we estimate the miscalibration bias $1 - \mu_i$ under the simplifying assumption that $E(\bar{q}) = \bar{q}$. We have 5 reported self-confidence levels in total (ranging from 0 to 100). We first ask subjects' self-confidence in succeeding each level during level 1, at the end of stage four ($q_i^{2,4}, q_i^{3,4}$). We repeat these questions just before level 2 starts ($q_i^{2,9}, q_i^{3,9}$)²⁰, so only for subjects who succeed level 1 and decide to continue. As success is a binary variable (either pass or fail the level) we cannot estimate the individual $1 - \mu_i$. We use then the mean success rates so we can estimate $1 - \mu_i$ for each treatment and levels, as follows:

- For low ability subjects, $1 - \mu_L = \frac{q - \bar{q}}{1 - \bar{q}}$ (from equation (1.1)), where \bar{q} is the observed mean success rate at each level and q the mean confidence level reported by the subjects. If low ability subjects tend to be overconfident then $1 - \mu_L > 0$ ($q > \bar{q}$), if they are under-confident $1 - \mu_L < 0$ ($q < \bar{q}$), and if they are accurate $1 - \mu_L = 0$ ($q = \bar{q}$).
- For high ability subjects, $1 - \mu_H = \frac{\bar{q} - q}{\bar{q}}$ (from equation (1.3)), if high ability subjects tend to be under-confident then $1 - \mu_H > 0$ ($q < \bar{q}$), if they are overconfident then $1 - \mu_H < 0$ ($q > \bar{q}$), and if they are accurate $1 - \mu_H = 0$ ($q = \bar{q}$).

19. There are 47,5% of high ability subjects in the GT, 56,1% in IT and 45.7% in the LT. A two-sided proportion test yields a p -value $_{GGvsGI} = 0.27$, p -value $_{GGvsGP} = 0.82$, and p -value $_{GGvsGI} = 0.18$.

20. At the end of stage 9.

Concerning low ability subjects (see table 1.6), we first observe that they are overconfident as they overestimate their probability of success ($1 - \mu > 0$ that is $q > \bar{q}$). Moreover, $1 - \mu_L$ is steadily increasing with the level of aspirations confirming the first part of assumption 1 (the miscalibration parameter increases with aspiration levels).

	GT	IT	LT		GT	IT	LT
$1 - \mu_L^{2,4}$	0,18	0,39	0,41	$1 - \mu_L^{3,4}$	0,29	0,34	0,46
$1 - \mu_L^{2,9} (/C1=1)$	- 0,06	0,23	0,14	$1 - \mu_L^{3,9}$	0,20	0,37	0,58

Note: $1 - \mu^{\ell,4}$ refers to the confidence reported at the end of stage 4, level 1 and $1 - \mu^{\ell,9}$ to the confidence asserted at the end of stage 9, level 1. C1=1 means that is it conditional on the decision to continue to the second level.

Table 1.6: Estimation of $1 - \mu$ for **low-ability** subjects

	GT	IT	LT		GT	IT	LT
$1 - \mu_H^{2,4}$	0,23	0,21	0,26	$1 - \mu_H^{3,4}$	na	na	na
$1 - \mu_H^{2,9} (/C1=1)$	0,29	0,33	0,29	$1 - \mu_H^{3,9}$	na	na	na

Note: $1 - \mu_4$ refers to the confidence reported at the end of stage 4, level 1 and $1 - \mu_9$ to the confidence reported at the end of stage 9, level 1. na: not available.

Table 1.7: Estimation of $1 - \mu$ for **high-ability** subjects

For high ability subjects (see table 1.7) the effect of aspirations is less clear. When looking at confidence for level 2 success ($1 - \mu_H^{2,4}, 1 - \mu_H^{2,9}$), values of $1 - \mu$ do not really differ across treatments, and subjects are equally underconfident ($1 - \mu > 0$ i.e. $q < \bar{q}$). This result corroborates the second part of assumption 1 (the miscalibration parameter varies little with the level of aspiration). At level 3 ($1 - \mu_H^{3,4}, 1 - \mu_H^{3,9}$), however, we cannot compare estimates of $1 - \mu$ because GT (low aspiration) subjects are slightly overconfident whereas IT subjects are underconfident and LT subjects are accurate. In this case, it is impossible to assume, as we did for building tables 5 and 6, that the ability-estimation bias is negligible. The importance of this remark will appear in the forthcoming discussion.

6.2. Aspiration upgrade and its effect on success

This section is dedicated to show any treatment differences on confidence and success. We make the hypothesis that the induced aspiration, later revised or not, will imply differences in confidence levels and hence on success.

6.2.1. An analysis of confidence

Figures 1.5 and 1.6 compare the success rate at level 2 and level 3 with the confidence level, by treatment (initial induced aspiration level) and ability. We can see that among low-ability subjects (figure 1.5), IT and LT subjects are over-confident in succeeding level 2 and especially level 3.

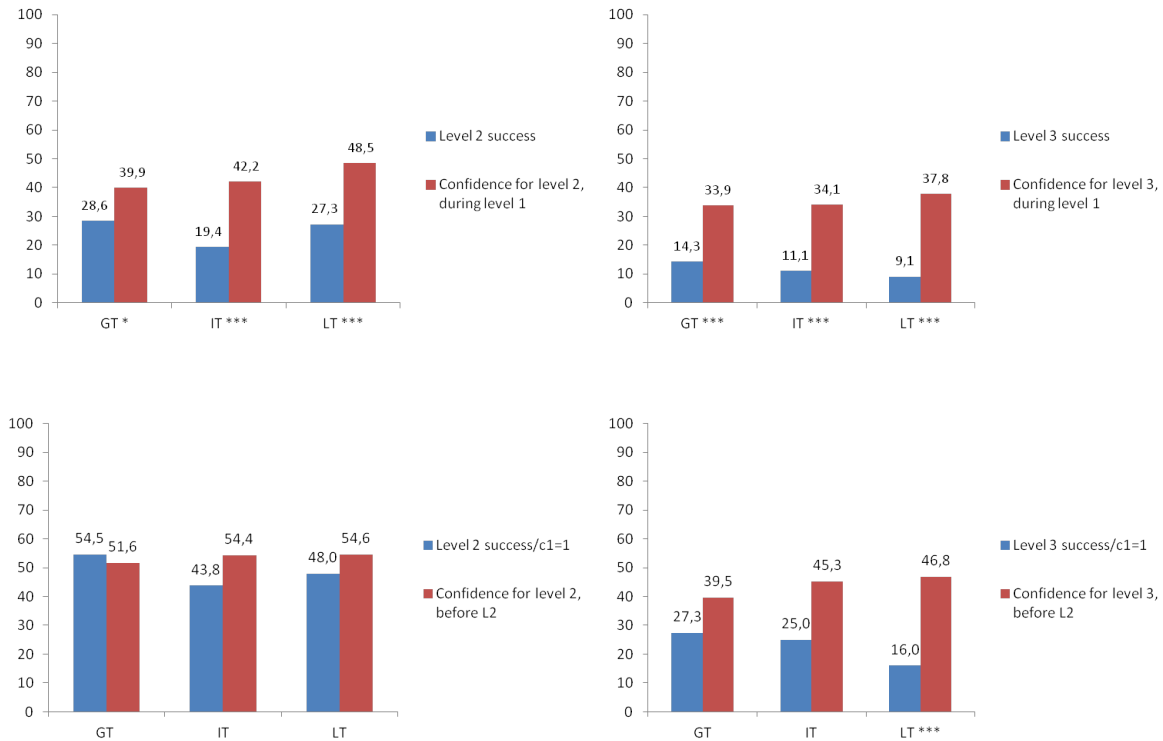


Figure 1.5: Confidence levels and success rate for **low ability** subjects (in %). Significance levels of two-tailed z-test: * 10% ** 5% *** 1%

Conversely, high-ability participants (figure 1.6) exhibit underconfidence in all treat-

ments at level 2. These two results taken together confirm the hard-easy and ability effects. However, a new result emerges from the strange behavior of high-ability IT participants at level 3 and, to a lesser degree, low-ability GT participants at level 2: these two groups exhibit underconfidence while they were expected to be overconfident. A look at figure 1.5 shows that this is not caused by an abnormally low level of confidence in comparison with other groups but to a relatively high rate of success. High-ability middle-class subjects who passed level 2 and low-ability lower-class subjects who passed level 1 both fulfilled their prior aspiration. We interpret their greater success at a higher level by the fact that they upgraded their aspiration level, thus gaining enough confidence to continue to a higher level and putting more effort in the task to succeed at this level if they are risk-averse (suggesting that ability and effort are complements in our experiment).

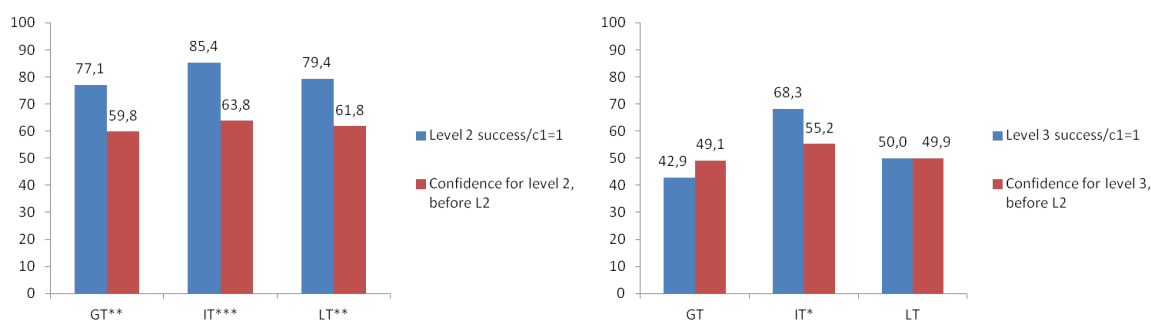


Figure 1.6: Confidence levels and success rate for high ability subjects (in %). Significance levels of two-tailed z-test: * 10% ** 5% *** 1%

The reference point's upgrading yields a new goal to achieve that seems to motivate them, allowing them to succeed better.

6.2.2. Aspiration, ability, and confidence effects on the decision to continue and success

In order to confirm this latter effect, we run probits on:

- The decision to continue to level 2 and level 2 success *conditional on succeeding*

level 1 for low-ability subjects to see if low-ability subjects from the GT subjects upgrade their aspiration by continuing more but not necessarily succeed then better. Table 1.8 reports this estimation.

- Both ability categories separately. We use two dependent dummy variables: *being relegated into the low group* (by failing to reach level 2) and *attaining middle or upper groups* (by succeeding to reach level 2 or beyond) *conditional on continuing to level 2* as level 2 corresponds to the initial level of aspiration for IT subjects. Table 1.9 and table 1.10 report these estimates²¹.

Inverse ability is captured by the average time needed to solve one anagram in the first four rounds, which is a rather exogenous measure. Ability 2 and Ability 4 refer to the second and fourth quartiles (the reference quartile is the first one which contains the 25% best participants)²².

Table 1.8 confirms our intuition. Indeed, among low-ability subjects who succeeded level 1, GT subjects, who fulfilled their prior and induced aspirations, have a higher probability to continue than IT subjects (see columns (1) and (2)), who did not already reach their initial aspiration levels. Even though they have low abilities, GT subjects are motivated by the positive surprise of having succeeded level 1. Nevertheless, they do not have a higher probability of success than the other treatments' subjects (cf. column (3) and (4)). Moreover, when controlling for individual characteristics, a higher confidence in level 2 success implies a higher probability of continuing to level 2 (column (2)). Results differ when considering success where a higher confidence in level

21. Two steps probits of success of level 1 and 2, conditional on continuing were run as well. They yield the same results for level 2 success. Concerning level 3, convergence problems arose because of a very few numbers of low ability subjects reaching level 3, as well as too many dichotomous variables in the model. However, for high ability subjects, results remain the same.

22. The same probit estimations were run with different ability variables: the number of total anagrams solved during level 1, the number of anagrams solved during the first four stages, the mean time of solving one anagrams during the whole level 1, the total time used in level 1: they all give qualitatively the same results. We choose to keep the mean time of solving one anagram during the first four stages because it represents, to our opinion, a more precise and exogenous ability variable.

Table 1.8: Probit on the decision to continue to L2 and L2 success (conditional on L1 success), on **low ability** subjects.

VARIABLES	Continue to L2 (1)	Continue to L2 (2)	L2 success (3)	L2 success (4)
Intermediate Treatment	-1.187*** (0.403)	-0.878** (0.442)	-0.376 (0.462)	-0.831 (0.565)
Loss Treatment	-0.514 (0.396)	-0.358 (0.423)	-0.173 (0.385)	-0.374 (0.450)
Ability 4	-0.642* (0.338)	-0.878** (0.362)	-0.140 (0.443)	-0.078 (0.523)
Confidence for L2 (during L1)	0.022 (0.014)	0.034** (0.016)		
Confidence for L3 (during L1)	0.003 (0.014)	0.003 (0.015)		
Confidence for L2 (before L2)			0.068** (0.029)	0.078** (0.034)
Confidence for L3 (before L2)			-0.049**	-0.042
Constant	0.345 (0.488)	0.816 (1.149)	-1.320 (0.861)	-4.334*** (1.603)
Controls	NO	YES	NO	YES
N	84	84	63	63

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: All variables except confidence levels (between 0 and 100) and age are dichotomous.

Control variables do not appear in the regressions but are included in the probit when controls is specified by "YES": Choosing "Hill"=1 of the track "Hill" has been chosen (which only concerns level 2 and 3) and its estimate is positive and significant at a 5% level for L2 success. Men=1 if the subject is a male participant and is non significant, stress=1 if the score is greater or equal to 4, risk aversion=1 if the subject chooses the sure payment and is negative and significant at a 5% level for the decision to continue, study levels=1 if the participant has three or more years of university education (or equivalent). French=1 if the mother tongue of the participants is French and is positive and significant at a 1% level for L2 success. Scrabble=1 if she plays scrabble regularly or occasionally. Participation=1 if she has already participated to an experiment and is negative and significant at a 10% level for L2 success. Order ab corresponds to the order of the curriculum subjects were presented. Paris=1 if the experiment is in Paris and =0 if it is in Montréal and is positive and significant at a 10% level for L2 success.

2 success significantly increases the probability of succeeding this level, a higher confidence in succeeding level 3 decreases this probability (the coefficient becomes however insignificant when control variables are included, see columns (3) and (4)). The same

estimations were run on high ability subjects but no effect was found except that being in the IT increases the probability of succeeding level 2 which corroborates the following analysis (at this stage of the game where difficulty is low for high ability subjects, aspiration effect is reduced by an ability effect).

Table 1.9: Probit on being relegated to the low group

VARIABLES	Low ability (1)	Low ability (2)	High ability (3)	High ability (4)
Intermediate Treatment	0.856** (0.363)	1.168*** (0.422)	-0.182 (0.326)	-0.408 (0.381)
Loss Treatment	0.451 (0.349)	0.488 (0.397)	-0.032 (0.340)	-0.219 (0.361)
Ability 2			1.089*** (0.283)	1.198*** (0.321)
Ability 4			(0.349) (0.397)	(0.340) (0.361)
Confidence for level 1	1.163*** (0.330)	1.720*** (0.414)	-0.012 (0.010)	-0.014 (0.011)
Confidence for L2 (during L1)	-0.010 (0.010)	-0.008 (0.011)	-0.018 (0.021)	-0.036 (0.024)
Confidence for L3 (during L1)	-0.017 (0.015)	-0.046** (0.020)	0.019 (0.015)	0.028 (0.018)
Constant	0.011 (0.010)	0.024* (0.014)	-0.176 (0.569)	1.551 (1.211)
Controls	0.826* (0.485)	2.056* (1.182)	NO	YES
N	NO	YES	NO	YES
R^2	122	122	121	121
	0.2	0.4	0.2	0.3

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: Controls' estimates: for the low performers, mother tongue estimates is negative with 5% sign., playing Scrabble is positive with $p=0.004$, ab order is negative with $p=0.035$, Paris is negative with $p=0.09$. For high performers, ab order is positive with $p=0.02$.

By first looking at what determines success or failure of low ability participants (columns (1) of table 1.9 and 1.10), we find that ability is the main determinant of

Table 1.10: Probit on attaining middle and upper groups

VARIABLES	Low ability	Low ability	High ability	High ability
Intermediate Treatment	-0.376 (0.462)	-0.788 (0.562)	0.368 (0.360)	1.094*** (0.419)
Loss Treatment	-0.173 (0.385)	-0.364 (0.456)	0.022 (0.370)	0.444 (0.425)
Ability 2			-0.960*** (0.310)	-1.199*** (0.374)
Ability 4	-0.140 (0.443)	0.049 (0.570)		
Confidence for L2 (before L2)	0.068** (0.029)	0.080** (0.032)	0.031** (0.015)	0.034* (0.019)
Confidence for L3 (before L2)	-0.049** (0.022)	-0.044* (0.025)	-0.031** (0.013)	-0.026 (0.017)
Constant	-1.320 (0.861)	-4.529*** (1.635)	0.999 (0.611)	-0.997 (1.347)
Controls	NO	YES	NO	YES
N	63	63	110	110
R^2	0.1	0.42	0.1	0.41

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Choosing "Hill" estimate is positive and significant with $p < 0.05$. Otherwise, the same control variables are used as in the previous probits (for the low performers: Men is positive and significant with $p = 0.02$, education is positive with $p = 0.023$, mother tongue is positive with $p = 0.006$, playing scrabble is negative with $p = 0.006$, Paris is positive with $p = 0.07$. For high performers, age is positive with $p = 0.09$ and order is negative with $p = 0.000$).

their failure. Confidence at level 2 prevents from being relegated to the lower group and helps rising to the middle and upper groups. This variable seems to capture effort whereas confidence at level 3, holding confidence at level 2 constant, would indicate excessive (over)confidence of low achievers as it exerts a significantly negative influence on success in table 1.9 (the same effect is thus found as in table 1.8). We can see on table 1.9 that belonging to the IT significantly increases their probability of failing prior to level 2, while the same effect was visible but insignificant on table 1.4.

This may result from a higher propensity to miscalibrate of low-ability individuals with aspirations (assumption 1). The number of observations shows that there are almost twice as many high performers than low performers that pass the second level. Only 11 high ability subjects fail level 1 or passed level 1 and quit the experiment. The only characteristic that seems to prevent them from failing is to belong to the first ability quartile rather than to the second (see column (2) of table 1.9). However their success is strongly driven by the fact of belonging to the IT (column (2) of table 1.10).

Hence, these estimations show how confidence can lead to success by stimulating effort but also how excessive overconfidence (indicating a strong propensity to miscalibrate) can lead to failure by discouraging effort.

This section suggests that individuals benefit from becoming more ambitious as they fulfil their initially moderate aspirations). When there is no performance surprise, the low-ability ones are harmed by the large miscalibration that results from aspirations exceeding too much their true ability.

7. Discussion

In this paper, we introduced confidence, aspirations and risk attitude into the economic analysis of education or career choices. Students and workers must decide to engage into increasingly difficult tasks with the hope of earning big money if they succeed and a lot less if they fail. Our analysis may also bring insights into the behavior of entrepreneurs and traders engaged in an escalation of risky ventures. Confidence and aspirations play a role in economic analysis if agents face uncertainty about their ability to succeed, a reasonable assumption in most situations of real life.

We conclude by using our theoretical approach to answer a basic controversial question: Is confidence a good or a bad thing when it deviates from the true frequency of success in a real-effort task? Our simple answer to this important question is that

confidence is productive when it stimulates effort and counter-productive when it discourages effort. Our model, resting on very robust behavioral assumptions, concludes that, under uncertainty about one's ability, confidence is often a good thing when it is caused by mild optimism and past experiences of good performance but it is always a bad thing when it is caused by a propensity to miscalibrate probabilities of success. Conversely, experiencing bad performance is usually a bad thing and reducing the propensity to miscalibrate is always a good thing.

Our experiment has provided several illustrations of these theoretical conclusions. For instance, we showed that higher aspirations can reduce miscalibration for high-ability (risk-averse) agents and increase it for low-ability agents. This prevents high-ability agents of the intermediate treatment from falling into the lower level but has the opposite effect on low-ability agents of the same class. We also found that early performance of low-ability subjects of the gain treatment (low induced aspirations), that is, passing level 1, raised their (initially minimal) ambition and stimulated their effort and performance relative to intermediate treatment.

The general framing of the model and experiment can be applied to sociological theories of education which look at the impact of social background on educational achievement. By a simple manipulation of endowments, we simulated three levels of aspiration that we matched with the three levels of difficulty introduced in our experiment. This enables us to interpret averages at the treatment and ability level as representative of social categories commonly considered in the sociology of education. However, in contrast with the typical description of a two-class society in which differences in economic opportunities govern intergenerational inequality²³, the adjunction of a middle class (intermediate treatment) elicits the surprising role played by the interaction of aspirations with ability. Indeed, if we make a simplifying parallelism with

23. After controlling for the direct effect of the cultural transmission by the parents, we still need to understand why well-endowed children are more encouraged to pursue their studies than others, given the same abilities, and if schooling institutions are more or less biased in selecting abilities.

our treatments and social classes, the middle class (IT) is characterized by an extreme intergenerational mobility, both upward for high-ability children and downward for low-ability children, which is less found in the lower (gain treatment) and upper classes (loss treatment).

Sociological theories (as Bourdieu and Passeron (1964)) describe the fact that individuals with a low level of aspirations are crowded out when the level of difficulty increases. They link confidence and aspirations such that individuals with a high level of aspiration are more confident than the ones with lower aspirations and hence succeed better. It turns out that our explanation of the hard-easy effect can accommodate this sociological conjecture by showing that $1 - \mu$ varies with the aspiration level A in systematic way.

8. Conclusion

Our behavioral model of confidence is a standard Bayesian subjective EU model but differs from existing papers in that it incorporates new psychological determinants. First, the most robust "anomaly" of the burgeoning literature on overconfidence: the hard-easy effect and the ability effect (Dunning-Kruger) that we consider as a corollary and second, aspirations. Our interpretation of the hard-easy effect is that people, on choosing whether to engage in a risk-taking activity, may miscalibrate their own probability of success even if they have an unbiased prior estimate of their own ability. This occurs if they are uncertain of the true probability of success because they can be misled by emotions triggered by their doubt. However, the direction of doubt is entirely different depending on whether their prior estimate led them to believe that they would fail or that they would succeed. Thus, we are led to consider two different types of miscalibration. Low-ability agents exhibit overconfidence because they are tempted by success although they should fail, whereas high-ability agents exhibit underconfidence because they are afraid to fail although they should normally succeed.

Our model predicts that the confidence bias is a function of two biases: an estimation bias on ability (ex ante), and a miscalibration bias that differs among the high and low ability agents. In this model, neither overconfidence nor underconfidence *per se* matter in terms of welfare because effort co-determines confidence and performance but does not determine their difference in any simple way. We further introduce aspirations in the model by assuming that people treat their desired goal as a reference no lower than the status quo which may shift upward after initial aspirations have been fulfilled. Differentials in aspirations interact with ability and risk attitude to determine the degree of miscalibration, and accordingly confidence, effort and performance. Indeed, effort will increase with a positive estimation bias but decrease with a higher miscalibration bias. Aspirations have a positive effect on risk averse high ability agents but should have little effect on low-ability subjects.

A. Appendix

A.1. Tables

	Montréal		Paris		Total
	AB	BA	AB	BA	
GT	16	15	19	30	80
IT	19	19	20	24	82
LT	20	21	20	20	81
Total	55	55	59	74	243

Table 1.11: Number of participants per session and in total

A.2. Figures

Niveau 1
Etape 1 sur 15

Décompte en mm:ss
06:32

1 2 9 10 11 12 13 14 15

6 anagrammes/étape 8 anagrammes 10 anagrammes 12 anagrammes

Niveau 1 Niveau 2 Niveau 3

iorevd devoir

nroec

nonred

rlprea

etmret

arvios

Nombre d'anagrammes résolues: 1 sur 54 Valider votre réponse

Figure 1.7: Experiment screen faced by subjects for level 1, step 1

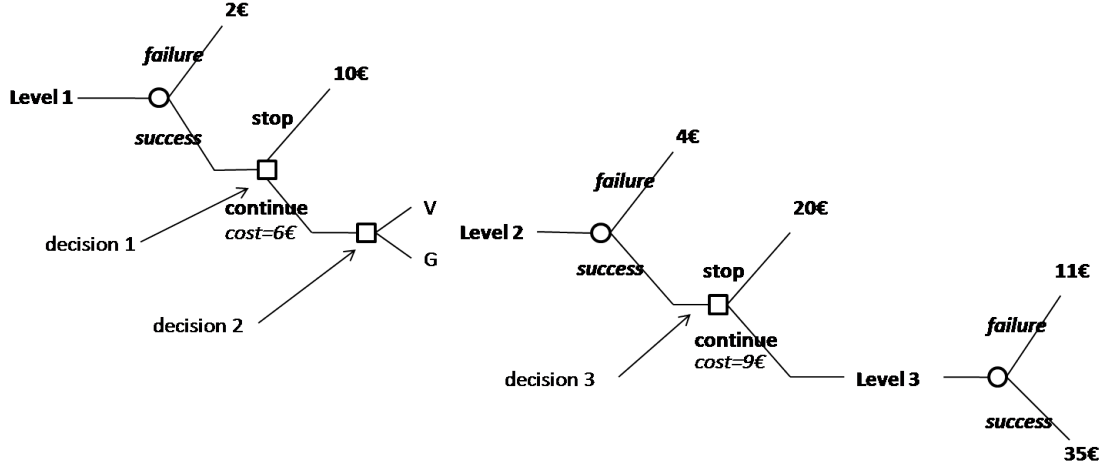


Figure 1.8: The decision process for a subject belonging to the Gain Treatment

A.3. Proofs

The derivative of the optimal level of effort with respect to the perceived level of abilities yields:

$$\frac{\partial e^*}{\partial E(k)} = \frac{\mu(A)\bar{q}''_{eE(k)}(e^*, E(k))W_l(A)}{c''(e^*) - \mu(A)\bar{q}''_e(e^*, E(k))W_l(A)} \quad (1.10)$$

The derivative of the optimal level of effort with respect to aspirations yields:

$$\frac{\partial e_i^*}{\partial A} = \frac{\bar{q}'_{e_i}(e_i^*, E(k_i))\left[\frac{\partial W_l(A)}{\partial A}\mu_i(A) + W_l(A)\frac{\partial \mu_i(A)}{\partial A}\right]}{c''(e_i^*) - \mu_i(A)\bar{q}''_{e_i}(e_i^*, E(k_i))W_l(A)} \quad (1.11)$$

With $\frac{\partial W_l(A)}{\partial A} = u'(y_{l,0} - A) - u'(y_l - A)$. This term depends then on agent's risk preferences: it is positive if the agent is risk averse ($u'' < 0$), negative otherwise.

The derivative of the optimal level of effort with respect to the calibration rate μ yields:

$$\frac{\partial e^*}{\partial \mu(A)} = \frac{\bar{q}'_e(e^*, k)W_l}{c''(e^*) - \mu(A)\bar{q}''_e(e^*, k)W_l} > 0 \quad (1.12)$$

A.4. Instructions of the experiment (translated from French)

NB: Specific instructions according to the treatment are detailed.

You are participating to an experiment in which we ask you to solve anagrams and to make decisions. Each participant makes his/her decision individually in front of his/her computer.

Before starting the experiment, you will have to answer a questionnaire about your age, occupation etc..., aimed to better know you. Answers to these question are very important for us and will be kept entirely anonymous.

For the Loss Treatment: **You are endowed with €35.** According to your decisions and performances you can loose part or all of this endowment. You will then receive at the end of the experiment €35 minus the loss you may deal with during the experiment.

For the Intermediate Treatment: **You are endowed with €20.** According to your decisions and performances you can loose part or all of this endowment, or winning €15 more.

To thank you for your participation, we will give you a participation fee of €5.

The experiment contains 15 steps combined in three levels.

The task:

At each step, you must **solve anagrams**. The principle is to makeup a word with the letters that are jumbled up and appear on the screen.

Example: You see "jrbnoou" on the screen. You must find and write "bonjour".

Please note:

- i. You must find a specific word, and not any word. Example: You must find "balle" from the letter "ablel". The word "label" will be then refused, even though it is correctly written. Do not be surprised if existing words are declined. In other words, there is only one correct answer for each anagram.
- ii. The words can be written with accents, but the anagrams and their solutions will be spelled : "emme" => "meme".
- iii. Solutions are not conjugated words, and cannot be applied to feminine or plural. You do not have to find any solutions such as "trouvas", "trouvait", "trouvées". If it is a verb, it will always be the infinitive form.
- iv. Capital letters are not accepted.

Many anagrams will be presented on the same screen and this set of anagrams will form a **step**. Many steps are grouped into one **level**. The experiment contains 3 levels. **To succeed on level, you must solve at least 2/3 of the total anagrams presented on all the steps forming this level.**

Level 1: Steps from 1 to 9

The first level groups **9 steps**. Each step consists in **6 anagrams** that can be solved. For each of the steps, the 6 anagrams will be presented at the same time, on the same screen. You have to type the solution words in the spaces provided for that. When a typed word is correct, the box near the word is checked.

You have **8 minutes maximum** to solve the anagrams displayed at **each step**. When you have no minute left, you are directly invited to the next screen.

You can skip to the next step without solving all the anagrams, and before the 8

minutes are up, by clicking on the "Submit your answer" box. For achieving Level 1 you must have solved **at least 36 anagrams** (out of 54) at the end of the ninth stage

For the Gain Treatment:

- At the end of level 1, if you have solved **less than 36 anagrams, the experiment is over**, and you win €2. You can leave your sit after obtaining permission from the experimenter by raising your hand. Your payoffs will be issued to you before you leave the room.
- If, at the end of level 1, you have solved **at least 36 anagrams**, you win €10. You can either choose **to stop there** and leave the experimental lab with your payoff, or **to continue the experiment** to level 2 by giving the money in.

For the Intermediate Treatment:

- At the end of level 1, if you have solved **less than 36 anagrams, the experiment is over**, and you and you loose €18. You can leave your sit after obtaining permission from the experimenter by raising your hand. Your payoffs will be issued to you before you leave the room.
- If, at the end of level 1, you have solved **at least 36 anagrams**, the amount of your loss is €10. You can either choose **to stop there** and leave the experimental lab with the remaining amount, or **to continue the experiment** to level 2 by giving the money in.

For the Loss Treatment:

- At the end of level 1, if you have solved **less than 36 anagrams, the experiment is over**, and you lose €33. You can leave your sit after obtaining permission from the experimenter by raising your hand. Your payoffs will be issued before you leave the room.

- If, at the end of level 1, you have solved at least 36 anagrams, the amount of your loss is €25 . You can either choose **to stop there** and leave the experimental lab with the remaining amount, or **to continue the experiment** to level 2 by giving the money in.

If you decide to continue the experiment, you must also choose between two options, coded A or B, for the remainder of the experiment (see attached diagram).

The number of anagrams solved in the rest of the experience depends on the option selected:

	A	B
Steps 10 to 12 (Level 2):	8 anagrams/step	10 anagrams/step
Steps 13 to 15 (Level 3):	12 anagrams/step	10 anagrams/step

This choice is final until the end of the experiment and cannot be changed at the end of level 2.

Level 2: Steps from 10 to 12

The second level consists of **3 steps**.

At each step, 8 anagrams if you chose Option A, 10 anagrams if you chose Option B, are presented.

You have 8 minutes per step. If the **8 minutes** are exceeded, you are invited to go directly to the next step.

You can skip to the next step without solving all the anagrams, however, it is required that you have solved **at least** 16 anagrams out of 24 if you chose Option A and 20 anagrams out of 30 if you have chosen Option B, at the end of the 12th stage for achieving Level 2.

For the Gain Treatment:

- At the end of level 2, if you have solved less than $2/3$ of the anagrams presented during the three stages, the experiment is over, and you win €4. You can leave your sit after obtaining permission from the experimenter by raising your hand. Your total payoffs will be issued before you leave the room.
- If, at the end of level 2, you have solved **at least $2/3$ anagrams**, you win €20. You can either choose **to stop there** and leave the experimental lab with your payoff, or **to continue the experiment** to level three by giving the money in.

For the Intermediate Treatment:

- At the end of level 2, if you have solved less than $2/3$ of the anagrams presented, **the experiment is over**, and you loose €16. You can leave your sit after obtaining permission from the experimenter by raising your hand. Your payoffs will be issued to you before you leave the room.
- If, at the end of level 2, you have solved **at least $2/3$ anagrams**, the amount of your loss is €0. You can either choose **to stop there** and leave the experimental lab with the remaining amount, or **to continue the experiment** to level three by giving the money in.

For the Loss Treatment:

- At the end of level 2, if you have solved less than $2/3$ of the anagrams presented, **the experiment is over**, and you loose €31. You can leave your sit after obtaining permission from the experimenter by raising your hand. Your payoffs will be issued to you before you leave the room.
- If, at the end of level 1, you have solved **at least $2/3$ anagrams**, the amount of your loss is €15. You can either choose **to stop there** and leave the experimental

lab with the remaining amount, or **to continue the experiment** to level three by giving the money in.

Level 3: Steps from 13 to 15

The third level consists of **3 steps**.

At each step, 12 anagrams if you chose Option A, 10 anagrams if you chose Option B, are presented.

You have **8 minutes** per step. If the 8 minutes are exceeded, you are invited to go directly to the next step.

You can skip to the next step without solving all the anagrams, however, it is required that you have solved **at least** 24 out of 36 anagrams if you chose Option A, 20 anagrams out of 30 if you have chosen Option B, at the end of the 15th step to complete the level 3 successfully.

For the Gain Treatment:

At the end of stage 15, the experiment is over. If you have solved less than $2/3$ of the anagrams presented during the three stages you win €11. If you have solved **at least $2/3$ of the anagrams**, you win €35.

For the Loss Treatment: At the end of stage 15, the experiment is over. If you have solved less than $2/3$ of the anagrams presented during the three steps, you lose €24. If you have solved at least $2/3$ of anagrams, the total amount of your loss is €0.

For the Intermediate Treatment: At the end of stage 15, the experiment is over. If you have solved less than $2/3$ of the anagrams presented during the three steps, you

loose are €9, subtracted to the €20 endowment. If you have solved at least 2/3 of anagrams, you get €15 more are to be added to the initial €20.

Further details

It is not possible for you to return to the previous step once you have clicked on "Submit your answer."

It is not possible for you to stop the experiment during a level.

You are not allowed to communicate during the experiment. If you have any questions during the experiment, raise your hand, someone will come to you individually.

It is very important that you understand these instructions. If you have any questions about these instructions, please raise your hand, someone will come to answer.

Thank you for your participation.

Payoffs' structure during the experiment (only one table was seen by the subjects, according to their treatment session):

Gain Treatment

	Level 1	Level 2	Level 3
Success	€10	€20	€35
Failure	€2	€4	€11

Intermediate Treatment

Initial endowment: €35

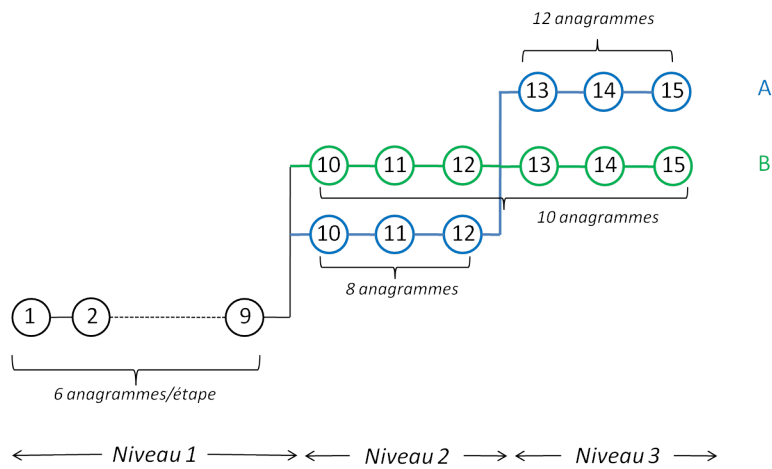
	Level 1	Level 2	Level 3
Success	-€25	-€15	-€0
Failure	-€33	-€31	-€24

Loss Treatment

Initial endowment: €20

	Level 1	Level 2	Level 3
Success	-€10	-€0	+€15
Failure	-€18	-€16	-€9

Experimental design:



Screen seen during step 1:

The screenshot shows the experimental interface during step 1. The screen displays the experimental design diagram, a timer (06:32), and a list of anagrams. The first anagram is "jrbnoou" and the second is "bonjour", which is marked as correct with a green checkmark. Below the list are five empty boxes labeled "Anagramme". At the bottom, it shows "Nombre d'anagrammes résolues: 1 sur 54" and a "Valider votre réponse" button.

A.5. Comprehension questionnaire for a GT subject in the A-B order

Please fill in the following questionnaire:

- i. In total, the experiment contains _____ levels. (Answer: 3)
- ii. The first level groups _____ stages. (Answer: 9)
- iii. Can you stop the experiment during the first level? Yes No (Answer: No)
- iv. For each step, the anagrams are shown on the same screen : Yes No
(Answer: Yes)
- v. Each stage lasts no more than _____ minutes. (Answer: 8)
- vi. During the first level, at each stage, you have been shown _____ anagrams.
(Answer: 6)
- vii. On the whole first level, at each stage, you must solve a total of _____ anagrams
to complete that level. (Answer: 36)
- viii. Assume you have chosen Option A at the end of level 1, how many anagrams do
you have to solve at least on level 2, to complete that level? _____
(Answer: 16)
- ix. At the end of level 2, can you change your mind on the choice of Option A or B
made at the end of level1? Yes No (Answer: No)
- x. Can you interact with other participants during the experiment? Yes No
(Answer: No)
- xi. What is the highest possible payoff you can earn? €_____ (Answer: €35)

**GENDER DIFFERENCES IN REACTION TO
FEEDBACK AND WILLINGNESS TO COMPETE**

This chapter is a joint work with Marie-Pierre Dagnies¹.

1. It has been submitted to the Journal of Economic Behavior & Organization.

1. Introduction

There are many possible explanations for the under-representation of women at the top of labor-market hierarchies, among which discrimination and the possibility that women may value the time spent with their children more. In this context, economists have recently become interested in the role played by gender differences in preferences (Croson and Gneezy, 2009).

Interest has in particular focussed on the gender gap in competitiveness. Men have often been found to have a greater taste for competition than women (Gneezy, Niederle and Rustichini, 2003, Niederle and Vesterlund, 2007, Datta Gupta, Poulsen and Villeval, 2012, Niederle and Vesterlund, 2011), whether this taste be measured by the selection decision into tournaments or by performance in a tournament imposed on subjects. There are however a number of exceptions. Gneezy, Leonard and List (2008) find that when choosing between piece-rate and tournament-incentive schemes, the gender gap in tournament selection is inverted in matrilineal societies (i.e. women select the tournament more often than do men). The gender gap in competitiveness also seems to be affected by whether the task is stereotypical-male or stereotypical-female, and by the level of pressure under which it is performed. In particular, the gender gap in tournament performance disappears when the task involves words rather than Maths (Shurchkov, 2012, Gunther et al., 2010) and under low-pressure conditions (Shurchkov, 2012), in which participants have more time than they need to perform the task.

It is sometimes argued that men are better-suited to hold prestigious positions than are women. We therefore focus here on a task for which men are known to self-select into competition more often than do women.

The gender gap in competitiveness has recently been found to disappear when participants are provided with information on their relative performance. Wozniak,

Hardbaugh and Mayr (2011) provide participants with a feedback telling them how participants in their session exactly performed in the piece rate. While there is a significant gender gap in tournament entry without feedback, it disappears when feedback is provided. Indeed, high-ability women choose more competitive compensation schemes and low-ability men choose less competitive compensation schemes with feedback than without it.

The goal of the present paper is twofold. We first evaluate how feedback on relative performance is perceived by men and women, that is how they update their beliefs following its reception. Second, we are interested in how men and women adjust to this information in terms of tournament entry decisions.

A recent literature has addressed the first point, i.e. how men and women differ in the way they react to the reception of performance feedback and how they subsequently update their beliefs. This is especially important as it may help rethink the way in which feedback is provided to employees in firms (giving precise information on past relative performance during annual reviews vs. setting goals for the future without focusing too much on past rankings) or to children at school.

Möbius et al. (2011) provide their subjects with noisy feedback via a simple binary signal for their performance being in the top 50%. They find that subjects update their beliefs about their IQ being in the top 50% less than Bayesian agents would in response to both positive and negative signals, and women update less than do men. They also show that subjects react more to positive than to negative information (and there is no gender difference in this respect). A number of papers (Möbius et al., 2011, Ertac, 2011, Grossman and Owens, 2011) find that individuals deviate from Bayesian beliefs more in self-relevant contexts (i.e. when they have to evaluate their own relative performance) than in self-irrelevant contexts (i.e. when they have to evaluate somebody else's relative performance, or update their beliefs about a neutral event). In terms of the deviation from Bayesian updating, we differ from the existing literature in that

our feedback is not noisy (contrary to Möbius et al. (2011) and Grossman and Owens (2011)) and subjects are asked to assess their beliefs over their relative performance in a competitive context (while Ertac (2011) uses a task remunerated under piece-rates).

Regarding our second point (how men's and women's actions react to performance feedback), Azmat and Iriberry (2010) find that providing feedback on relative performance to high-school students improves their grades by 5% regardless of where they are in the distribution. In the laboratory, this feedback-performance effect is only found for men (Azmat and Iriberry, 2012).

In our set-up, subjects have to decide twice whether to enter a tournament. The first time, the subject knows the opponent will be randomly-selected from all the other participants in their session, and will therefore be of totally unknown ability. After the participants make this first decision and perform the task, they receive a binary feedback telling them whether their performance was above or below the median in their session². We have two treatments allowing us to manipulate the degree of competition our subjects face. To our knowledge, our paper is the first one to directly manipulate the level of competition and study how this affects competitiveness while carefully monitoring relative confidence after the provision of a feedback. In the *ability group* treatment, they then decide a second time whether to enter a competition knowing that their opponent will be randomly selected among participants belonging to the same performance group as their own. In the *Repetition* treatment, they decide a second time whether to enter a competition with an opponent of totally unknown performance level (i.e. again randomly chosen from all the participants in the session). We are thus able to control for order effects. We make sure to elicit beliefs both before and after subjects receive their performance feedback. We can then see how beliefs and their updating affect the tournament entry decision.

Our first result is that subjects update their beliefs following performance feedback

2. Contrary to Wozniak, Hardbaugh and Mayr (2011), who provide an exact performance feedback based on the piece-rate, we provide a binary performance feedback based on the tournament.

more than would a Bayesian agent. Both men and women are more pessimistic than a Bayesian agent following negative feedback; we find the opposite effect after positive feedback. Both of these effects are stronger for women than for men. Our paper also shows that low-performing participants adapt their tournament entry decision to the ability level of the competition, while high-performing participants do not.

Men and women do not react to the feedback in terms of competitive entry in the same way. While women are especially sensitive to information on their own performance, men react more strongly to the level of their competitors. Feedback does not therefore seem to be processed in the same way by men and women. While men, and especially low-performing men, seem to take into account the possibility that their performance will improve over time, this is not the case for women. In other words, low-performing women consider their performance level *per se* while men think there is room for improvement.

Regarding the efficiency of choices, men mostly enter in about the proportion suggested by payoff maximization, but not enough women choose the tournament. This can be explained by women giving too much weight to the negative feedback they receive, and not taking into account the fact that their performance can (and does) improve over time.

The remainder of the paper is organized as follows. Section 2 describes the experimental design, section 3 formulates the different research questions and the analytical method. Section 4 reports the results on beliefs, treatment effects and establishes a welfare analysis. Section 5 then discusses our results and concludes.

2. Experimental design

We use a real effort task consisting in solving as many sums of five two-digit numbers as possible (Niederle and Vesterlund, 2007) within 5 minutes. There are a total of six

sequential steps³, one of which is randomly selected at the end of the experiment to determine the payoff, as well as incentivized belief-assessment questions. The fact that only one randomly chosen step determines the payoff eliminates any hedging opportunity. The *Repetition* and *ability group* sessions differ only in steps 4 and 4 prime (see below)⁴. The same number of men and women take part in each session, and each session includes at most 20 participants.

Step 1: Piece-rate (PR) remuneration scheme. Subjects have 5 minutes to solve as many sums as they can, and earn €0.50 per correct sum. This step allows to obtain a performance level for each subject under piece-rate.

Step 2: Standard tournament (ST). Subjects have 5 minutes to solve as many sums as they can. They are randomly paired with another player in their session (whose gender is unknown to them). If step 2 is randomly chosen for remuneration, the winner in each pair (with the best step-2 performance) earns €1 per correct sum, while the loser receives nothing. This step indicates the performance level of subjects under a tournament incentive. We furthermore use this step to obtain the performance one will have to compete against in future tournaments.

First round of belief elicitation: After the second step, the participants have to evaluate the probabilities that their step-2 performance belongs to each of the four performance quartiles. The sum of these 4 probabilities (in %) is equal to 100⁵. They answer four questions corresponding to the four quartiles. For example, the question for the fourth quartile was: "What is, according to you, the

3. The fact that the tasks are completed sequentially may have an effect on subjects' decisions to compete, as learning and fatigue may affect these decisions. However, as this paper focuses on gender effects and differences across treatments, the relevant question is whether men and women, or participants randomly assigned to different treatments, are affected in a different way by the tasks being sequential; this is, in our opinion, unlikely.

4. The instructions read to the subjects can be found in the Appendix.

5. The subject is asked to enter four beliefs: one for each quartile. If the sum of these four beliefs is not exactly equal to 100, the participant cannot go on to the next step and has to enter new beliefs adding up to 100.

probability in % that your step-2 performance belong to the 4th quartile (being in the 25% best performers)?"

To incentivize the answers, we use a confidence rule (Möbius et al., 2011, Hollard, Massoni and Vergnaud, 2010): for each of the four answers, the computer randomly picks a number y between 0 and 100. Let x_i be the subject's answer for quartile i ($i = 1, 2, 3, 4$).

- If $x_i > y$, the subject earns €1 if her score belongs to the i^{th} quartile, nothing otherwise.
- If $x_i < y$, the subject earns €1 with $y\%$ probability⁶.

Step 3 (hereafter Choice 1): Before solving the sums, subjects have to choose between the Piece Rate (PR) and Standard Tournament (ST) remuneration schemes. Those who choose PR receive €0.50 per correct sum if step 3 is picked at the end of the experiment. If a subject chooses the tournament, she is randomly paired with another subject and wins the tournament (which pays €1 per correct sum) if her step-3 performance is greater than her opponent's step-2 performance⁷. This step allows to replicate previous results in order to check any gender gap in competitive entry. It also gives us a benchmark for the gender gap in competitive entry in our experiment.

6. This rule can seem complicated. Our main goal was to get the beliefs on the quartiles before and after the feedback in order to compute the bayesian beliefs and be able to tell whether participants were updating their beliefs in a bayesian way. We had the choice between the quadratic scoring rule (QSR) and the method we ended up choosing. It appeared to us that both methods are complicated. The argument in favor of our chosen mechanism was that it is always truth-inducing while QSR only is for risk-neutral subjects. During each session, we carefully explained the rule and illustrated it with various example to make sure subjects understand it. We told participants that even though they may not understand all the elements of this rule, it was done to make sure they tell the truth about their beliefs. Hence, by not doing so they would be worse-off.

7. As the randomly-chosen opponent may not have chosen the tournament, the step-3 performance of the participant is compared to the step-2 performance of the opponent, when he was performing in a tournament. This way, the decision to enter the tournament is not affected by beliefs about whether the opponent is going to enter. In addition, it allows us to rule out the possibility that a participant may not enter because she does not want to inflict a loss on her opponent.

Step 3 prime (hereafter Choice 1 prime): Participants have to choose between *submitting* their step-1 performance to PR or ST. They do not have to solve sums at this step. The payoffs depend only on their step-1 performance. If a subject chooses to submit this performance to the tournament, she is randomly paired with another participant and earns €1 per correct sum if her step-1 performance is greater than her opponent's. If she chooses PR, the remuneration is the same as in step 1. Step 3 prime is identical to Step 3 (in both cases the tournament is a riskier choice implying more ambiguity and only giving the information whether she beat her opponent at the end of the experiment) except for the fact that it does not involve a future performance. In particular, the participant who chooses to submit her past performance to the tournament does not have to perform under the pressure of competition. In other words, Step 3 allows for improvement in performance (since the Step 3 performance of a participant will be compared to the Step 2 performance of her opponent) while the Step 3 prime tournament does not (Step 1 performances of both opponents are compared in Step 3 prime tournament).

Feedback: Each participant receives feedback on their step-2 performance, telling them whether it was above or below the median.

Second round of belief elicitation: This second round allows us to analyze how subjects update their beliefs after receiving an ability signal. They have to re-estimate the probabilities that their step-2 performance was in the two possible quartiles consistent with their feedback (the fourth and third quartiles for performers above the median, the second and first quartiles for performers below the median). We use the same incentive rule as in the first round, for both of the elicited beliefs.

Step 4 *Repetition* or *Ability Group* (hereafter Choice 2): 5 minutes of sums

- In *Repetition* sessions, step 4 is exactly the same as step 3 (choice between PR or ST, the remuneration rule stays the same). We use this Step as a control for the pure effect of feedback on competitive entry when nothing else changes.
- In *Ability Group* sessions, subjects have to again choose between a piece rate and a tournament. If the piece rate is chosen, the subject earns €0.50 per correct sum. But if she chooses the tournament, she is randomly matched to another participant who belongs to the same ability group. That is, if her step-2 performance was below (above) the median she is paired with someone whose step-2 performance was below (above) the median as well. We call this the "Ability Group Tournament" (AT) (see Table 2.1). A subject wins if her step-4 score is greater than her opponent's step-2 score, when the opponent is in the same ability group. In this case she wins €1 per correct sum, nothing otherwise. This Step gives us our main treatment effect, that is whether participants adapt their competitive entry to the level of the tournament.

Repetition (control)	Ability Group (treatment)
Choice between PR and Tournament	Choice between PR and Ability Group Tournament.

Table 2.1: *Repetition* (control) and *Ability Group* (treatment) at step 4

Step 4 prime (hereafter Choice 2 prime): Participants have to choose between *submitting* their step-1 performance to a piece rate or a tournament. In *Repetition* sessions, step 4 prime has the same features as step 3 prime. In *Ability Group* sessions, the choice of competition leads to an ability group tournament such that both of the step-1 performances compared belong to the same ability group. The remuneration rule is the same as in step 3 prime. This step has the same purpose for Step 4 as Step 3 prime has for Step 3.

At the end of each step, all participants are told their absolute performance, that is the number of sums they solved. At the end of the experiment, a screen summarizes the earnings from each step, and whether a tournament was won or lost in the case it was chosen by the participant. One step is randomly chosen to determine part of their payoffs, in addition to the belief-assessment questions. Figure 2.1 summarizes all the steps of the experiment.

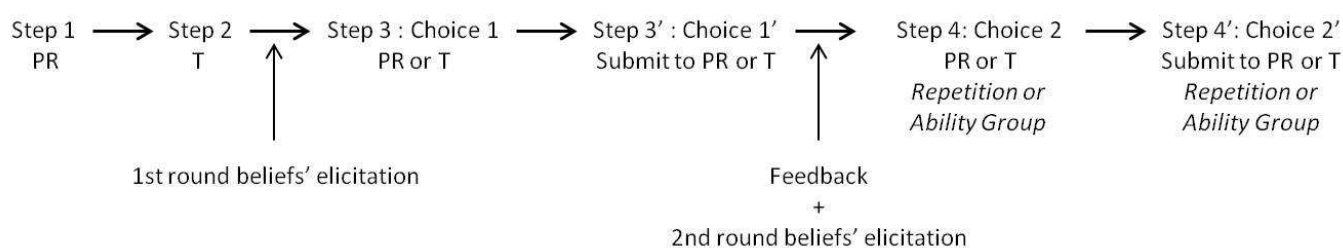


Figure 2.1: Experiment summary

While studying the effect of performance feedback on competitive entry has already been done (Cason, Masters and Sheremeta, 2010, Wozniak, Hardbaugh and Mayr, 2011), our paper is, to our knowledge, the first attempt to directly manipulate the level of competition one involves in. It allows us to study the combined effect of feedback and competition level while carefully monitoring beliefs about relative performance.

The experiment was run at the "*Laboratoire d'Economie Expérimentale de Paris*" (LEEP) between February and April 2011. The same number of men and women took part in each session. Respectively 112 subjects (56 men and 56 women) and 116 subjects (58 women and 58 men) participated in the *Repetition* and *Ability Group* sessions. One step was randomly chosen at the end of the experiment to be paid in addition to the belief-assessment questions and a €7 show-up fee. Participants earned €15.30 on average (see Table 2.7 in the Appendix for detailed sample characteristics).

3. Research questions and description of methods

We here state our research questions and set out the strategy used to answer them.

Our first research question is to know how confident men and women are over their relative performance level to start with, and how they update their beliefs following feedback. The strategy we use in Subsection 4.1 to study the belief-updating process consists of two steps. We first calculate a Bayesian benchmark from participants' priors about their performance being above the median (elicited before participants received any information) and their subsequent feedback. We then compare this benchmark to the beliefs elicited after participants received their feedback. This allows us to see whether our participants are Bayesian updaters, whether they update in the same way following positive and negative feedback, and whether men and women update their beliefs similarly.

Subsection 4.2 appeals to the diff-in-diff method to see whether our participants adapt their tournament entry decision to the information received about both their own performance and that of the opponent they will face. The diff-in-diff method seems appropriate in our case as we are interested in the change in competitiveness when manipulating the level of competition and providing feedback. In our *Ability Group* treatment, we are mostly interested in men and women's competitiveness after the manipulation of the level of competition (Step 4). However, this measure of competitiveness only makes sense in comparison with a "benchmark level of competition" which is measured in Step 3. Furthermore, the difference in competitiveness between Step 3 and Step 4 could be due to order effects, the effect of the performance feedback which is provided between Steps 3 and 4 and, finally, to the fact that the Step 4 tournament (contrary to the Step 3 tournament) opposes participants of the same performance group. This is why we use the *Repetition* treatment as a control, to see what the competitiveness difference between Steps 3 and 4 is when the choices

offered in both of these steps are exactly the same and where the only things which can explain a change in choice between the steps are order effects and the effect of feedback. The diff-in-diff method therefore allows us to isolate the effect of the *Ability Group* treatment *per se*.

As feedback depends on whether the participant's performance was above or below the session median, we calculate the diff-in diff estimators separately for these two groups. We ask whether the increase in the percentage of participants choosing tournament entry differs between choice 1 (before the feedback) and choice 2 (after the feedback) in the *Repetition* treatment (where the opponent is randomly-chosen from all session participants in both choices 1 and 2) and in the *Ability Group* treatment (where the opponent in choice 2 is randomly chosen from session participants in the same ability group, i.e either above or below the median).

In Choice 2, participants are matched to an opponent of unknown ability in *Repetition* and to a low-performing opponent in *Ability Group*. Low-performing participants⁸ who adapt entry to the level of the tournament should then increase entry more between Choices 1 and 2 in the *Ability Group* than in the *Repetition* treatments. On the contrary, high-performing participants should increase tournament entry more between Choices 1 and 2 in the *Repetition* than in the *Ability Group* treatment.

In Subsection 4.3 we link the belief-updating process to tournament entry. To do so, we construct the variable "beliefwin" to proxy the subject's beliefs regarding her chances tournament success. We then introduce "beliefwin" into the tournament-entry regressions to see whether it helps to explain individual choices.

Finally, Subsection 4.4 analyzes the welfare implications of the behaviors we observe. We calculate the expected payoffs from entering each tournament by drawing a great number of times without replacement from the set of the performances of the po-

8. Low-performing participants are those whose Step-2 performance was below the median Step-2 performance in their session, and who therefore receive the "below median" feedback after the first round of confidence-assessment questions.

tential opponents our participants face (e.g. participants whose performance is below the median for the "low-performing *Ability Group*" tournament). This shows the probability that a participant with a given performance level would win the tournament, and thus the expected payoff from tournament entry. We also compute the expected payoff from choosing the piece rate for each performance level. We can therefore determine the optimal choice, in expectation, at each performance level. The last step is to compare the percentage of participants who are expected to benefit from tournament entry to that who actually choose to do so.

4. Results

4.1. Changes in performance and confidence assessments

We start by analyzing how performance alters between steps 1 and 2, that is when the remuneration scheme changes from being piece rate to tournament.

Table 2.2 shows the average number of correctly-solved sums by gender in step 1 (piece-rate) and step 2 (tournament). Participants perform significantly better in the tournament than in the piece-rate. This suggests that the remuneration scheme affects performance, even though learning may also play a role here. We do not find any performance differences between men and women.

	Step 1 Performance (PR)	Step 2 Performance (T)	Diff
Men	7.8	9.2	$p=0.00$
Women	7.5	8.9	$p=0.00$
Diff	$p=0.48$	$p=0.46$	

Table 2.2: Average number of correctly-solved sums at step 1 and step 2 (the p-values correspond to two-tailed t-tests).

We pick up participants' confidence in their chances of tournament success via a question on their beliefs about their relative performance; this is asked both before

and after they receive feedback telling them whether their performance was above or below the median. Before receiving feedback, participants had to state their beliefs that their Step-2 performance fell in each of the four performance quartiles; after receiving feedback, they had to assess their beliefs that their performance fell in each of the two quartiles they could still appear in (as they now know that their performance was either above or below the median).

Before receiving feedback, low-performing men were not significantly more confident than low-performing women (thinking that their Step-2 performance was respectively 56.7% and 54.6% likely to be above the median Step-2 performance of their session): the two-sided Mann-Whitney test yields $p=0.69$. However, after feedback that their performance was below the median, low-performing men are more confident than low-performing women regarding their chances of belonging to the second as opposed to the first (worst) quartile (low-performing men and women respectively think they are 67.3% and 57.5% likely to belong to the second quartile: the two-sided Mann-Whitney test yields $p<0.01$).

Before receiving feedback, high-performing men were more confident than high-performing women of being better than the median participant. High-performing men and women respectively think that they have a 75.6% and 62.4% chance of being above the median (this difference is significant in a two-sided Mann-Whitney test with $p<0.01$). Once they learn that their performance is above the median, men are still more confident than women, but to a lesser degree: high-performing men and women believe their performance has respectively a 60.6% and 51.4% ($p=0.03$) chance of belonging to the 4th (top) quartile.

The above results suggest that men and women do not react in the same way to performance feedback, with women seeming to adjust more strongly to this feedback than do men. To further our investigation, we calculate for each subject the beliefs she would have held during the second round of confidence-assessment questions (i.e. after

performance feedback) were she to have updated her first-round beliefs in a Bayesian way. We later refer to these beliefs as "Bayesian beliefs". They are calculated as follows.

We denote by b_{1_i} the first-round elicited beliefs about belonging to the i^{th} quartile ($i = 1, \dots, 4$). We differentiate Bayesian beliefs for low performers ($Bayes_{low}$) and high performers ($Bayes_{high}$) such that:

- $Bayes_{low} = 100 * \frac{b_{1_2}}{b_{1_2} + b_{1_1}}$
- $Bayes_{high} = 100 * \frac{b_{1_4}}{b_{1_4} + b_{1_3}}$

We then compare actual second-round beliefs to Bayesian beliefs. Figure 2.2 displays the Bayesian beliefs compared to the actual updated beliefs after receiving the feedback, for low and high-performing participants, broken by gender. If our subjects were Bayesian updaters, beliefs should be situated on the 45° line that is where Bayesian beliefs are equal to the true beliefs. However, we can see that overall, most of the low-performing participants' beliefs are situated below the 45° line, that is they update more pessimistically than a Bayesian agent would. The opposite result is found for high-performing participants. More precisely, both men and women overreact to the feedback received, but women more so. High-performing women's beliefs are significantly more optimistic than Bayesian beliefs (a Wilcoxon signed-rank test yields $p < 0.0001$); this is also the case for high-performing men ($p < 0.01$). However, high-performing women update their beliefs significantly more optimistically than do high-performing men (a Mann-Whitney test yields $p = 0.04$).

Concerning low-performing participants, women's beliefs are significantly more pessimistic than Bayesian beliefs (a Wilcoxon signed-rank test yields $p < 0.0001$) while this holds to a lesser extent for men ($p = 0.04$). Furthermore, low-performing women update significantly more pessimistically than do their male counterparts ($p = 0.04$).

Result 1: *While both men and women overreact to the feedback they receive, women*

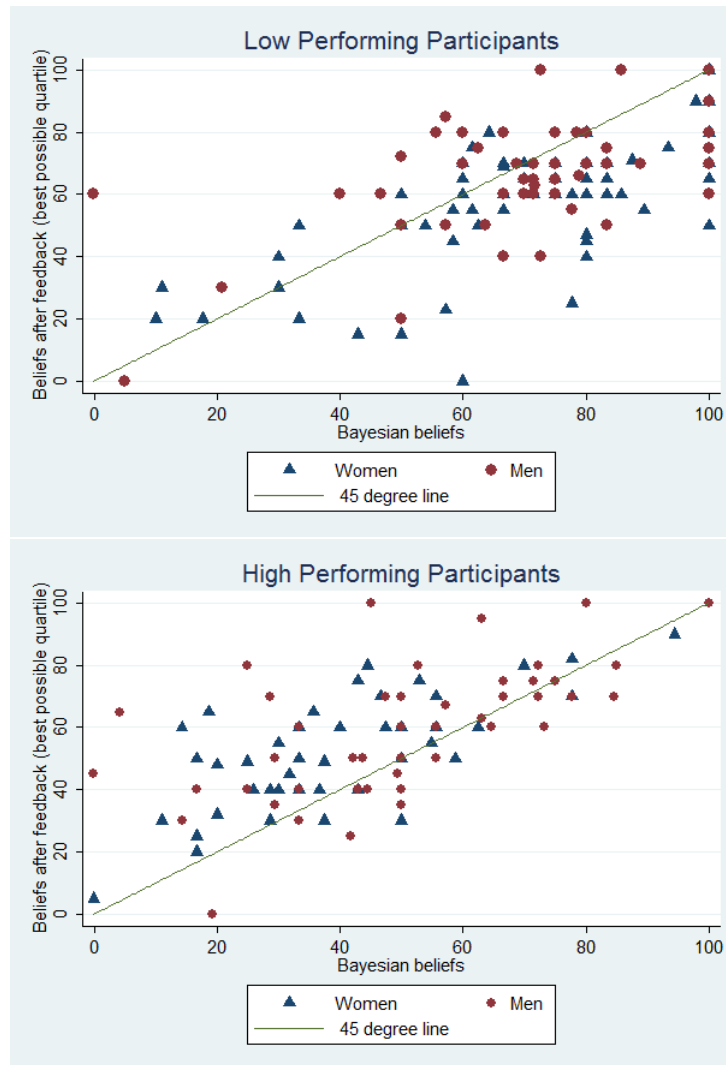


Figure 2.2: Beliefs update compared to Bayesian beliefs

do so much more than do men.

After being informed that their performance is below the median, low-performing women are too likely to think (compared to a Bayesian updater) that their performance is in the lowest quartile. On the other hand, the good news that high-performing women receive makes them too likely to think that they are in the top quartile. This first result suggests that there is no conservatism behavior (that is updating too little in response

to both positive and negative signals). It differs from Möbius et al. (2011) who use a noisy signal and find that subjects are conservative. Notice that the fact that our signal is not noisy (contrary to Möbius et al. (2011) and Grossman and Owens (2011)) does not give the chance to subjects of believing the information they receive is wrong. Ertac (2011)'s first finding is that updating is on average pessimistic especially following a negative feedback. This is in line with our result that subjects overreact to negative information. Secondly, the author finds that women tend to interpret positive feedback more conservatively than men. Our result is in contrast with this one. Two design features differences may explain this. First, Ertac (2011) asks her subjects their beliefs about belonging to three unequal performance categories (top 20%, middle 60%, top 20%) when we use quartiles and second, our subjects are asked to assess their beliefs over their relative performance in a competitive context while Ertac (2011) uses a task remunerated under piece-rates.

4.2. *Difference-in-difference analysis*

4.2.1. **On the tournament entry decision**

The answer to our main question can be seen by looking at whether the change in the tournament entry decision between Choices 1 and 2 is different for participants in the *Repetition* and *Ability Group* treatments. In both treatments, participants receive feedback between Choices 1 and 2 for whether their performance was above or below the median. However, only in the *Ability Group* does the opponent belong to the same performance group as the subject. If low-performing participants adjust their tournament entry decisions to the level of the competition, we expect this adjustment between Choices 1 (step 3) and 2 (step 4) to be greater in the *Ability Group* (where the level of the competition is lower in the choice 1-ability tournament than in the choice 2-standard tournament) than under *Repetition* (where it remains the same). The opposite should hold for high-performing participants. Figures 2.3 and 2.4 shows

the tournament entry rates of women and men, by ability level in the *Repetition* and *Ability Group* treatments.

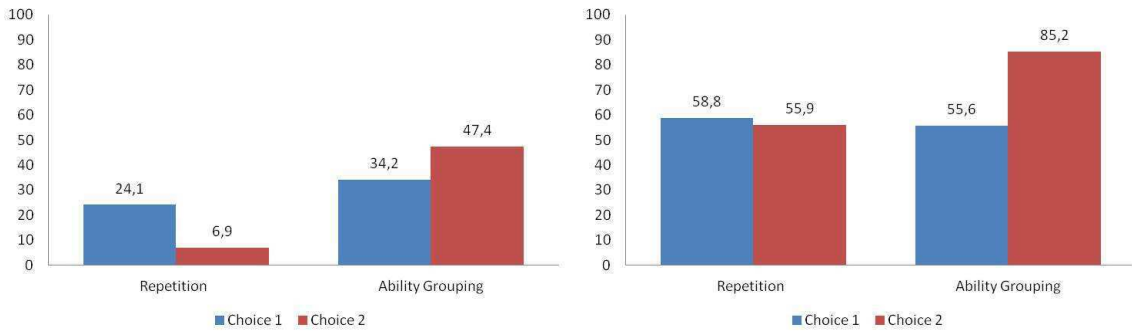


Figure 2.3. Proportion of low-performing women (left) and men (right) choosing tournament entry in Choice 1 and Choice 2.

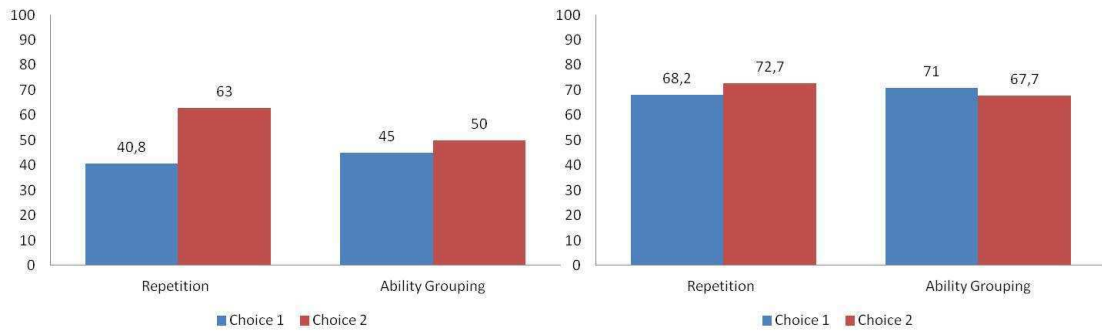


Figure 2.4. Proportion of high-performing women (left) and men (right) choosing tournament entry in Choice 1 and Choice 2.

In order to verify if the differences between tournament entry rates, before and after feedback are different according to the treatment we compute the diff-in-diff estimators for Choice 1 vs. Choice 2 tournaments and *Repetition* vs. *Ability Group* by estimating the following equation:

$$Tourn_{it} = \beta_0 + \beta_1 AbGr_{it} + \beta_2 Choice2_{it} + \beta_3 AbGr * Choice2_{it} + \epsilon_{it} \quad (2.1)$$

Regarding equation 2.1, *Tourn* is a dummy variable equal to 1 if subject *i* entered

the tournament. *AbGr* is the dummy treatment variable for the subject taking part in one of the *Ability Group* treatments. The diff-in-diff estimate is thus given by the coefficient β_3 on the interaction term *AbGr*Choice2*. These estimates are positive and significant for both low-performing women and men: the coefficients are respectively 0.33 ($p=0.02$) and 0.30 ($p=0.04$), indicating a treatment effect for those subjects, in that they adapt their entry decision to the competition level. Regarding high-performing participants, the diff-in-diff estimators are negative (respectively -0.08 and -0.17 for high-performing women and men) but not significant (respectively $p=0.23$ and $p=0.54$). High-performing participants do not therefore increase tournament entry more under *Repetition* (where the level of the competition remains the same) than in the *Ability Group* (where it is higher).

Result 2: *Low-performing participants adapt their choice to enter a tournament to the level of the competition, while high-performing participants do not.*

This new finding is consistent with the reference group neglect phenomenon found by Camerer and Lovo (1999). By selecting subject on their abilities, previous to the experiment, they actually found that these subjects have a higher propensity to enter a market even though it is not always optimal, compared to subjects who did not go through the selection process. Our result is however more subtle, showing that it only concerns subjects having a positive feedback: high ability participants neglect the level of the competition they will face.

We could be tempted to conclude from these first results that men and women react similarly to the level of the competition they face. However, finding that low-performing participants adjust their tournament entry decision to the level of the competition could reflect two different phenomena. Subjects could be reacting either to the feedback or the level of their opponent. Participants would be reacting to feedback if, for instance, the receipt of negative feedback reduced their willingness to compete in *Repetition* but

there was no difference in tournament entry between Choices 1 and 2 in the *Ability Group*, where they know that their opponent is also of low ability. On the other hand, were the receipt of negative feedback not to change the tournament entry decision in *Repetition*, but to do so in the *Ability Group*, participants enter more in the Choice-2 than in the Choice-1 tournament, then we would conclude that participants react more to the level of the competition *per se*.

We compute the diff-in-diff estimators for Choice-1 vs. Choice-2 tournaments and below median vs. above median for each treatment, by gender:

$$Tourn_{it} = \gamma_0 + \gamma_1 AboveMedian_{it} + \gamma_2 Choice2_{it} + \gamma_3 AboveMedian * Choice2_{it} + \epsilon_{it} \quad (2.2)$$

For the *Repetition* group, regarding equation 2.2, the estimate of the coefficient on the interaction term of *AboveMedian* and *Choice2* (which corresponds to the diff-in-diff estimate) for women is positive and significant (the coefficient is 0.39, with $p=0.01$), so that women increase tournament entry between Choice 1 (Step 3) and Choice 2 (Step 4) more following positive than negative feedback. The nature of the feedback therefore seems to have considerable impact on women's competitive decisions. This is not the case for men (the coefficient is 0.07, $p=0.57$), where tournament entry decisions are not affected by the nature of the feedback received.

For the *Ability Group*, men respond differently according to whether they are above or below the median (the coefficient is -0.33, $p=0.01$), so that they react more to what the feedback implies for the level of the competition they will face rather than the information on their own performance level. In this same group, the insignificant coefficient ($p=0.56$) for women underlines that they place more importance on the personal information than on the implication for their opponents' ability.

Result 3: *While women react mainly to feedback on their own performance level, men respond more to the level of their opponent.*

This is to the best of our knowledge a novel finding. It should be remembered that those entering a tournament at a given step will win if their performance at this step is greater than the Step-2 performance of their opponent. As such, if performance is expected to increase from one step to another, the probability of tournament success should be greater in later steps. The fact that men do not react to feedback in the *Repetition* treatment (i.e. when it only provides information about their own relative performance at Step 2) could mean that they expect their performance to improve over time (and more so when their performance is low to begin with). Women, on the other hand, seem to take the feedback at its face value.

4.2.2. On the tournament submission decision

The decision to submit to a tournament is used to control for any effect of risk (ambiguity and feedback) aversion⁹, since the latter should affect both submission and entry decisions equally. The decision to submit one's performance to a tournament is very similar to deciding whether to enter a tournament, except that the subject does not have to actually perform the task in the first case. As such, while overconfidence and risk or ambiguity aversion can play a role in the submission decision, the fear of choking under competitive pressure, and any beliefs that performance will improve over time cannot play any role here¹⁰.

We find no treatment effect for low-performing participants regarding the decision to submit one's past performance to the piece-rate or the tournament, contrary to the decision to actually enter the competition: for low-performing men and low-performing women, the diff-in-diff estimates (Choice 1 vs. Choice 2 and *Repetition* vs. *Ability*

9. We decided to use the decision to submit to control for risk aversion for two reasons. Firstly, it allows us to have a risk aversion measure in the same context than the main task. Secondly, we obtain two separate measures of risk-aversion, one in the context of Step 3 and one in Step 4 where feedback as been provided and when the level of the competition might change.

10. Whenever a participant submits her past performance to a tournament, she wins if her Step-1 performance is greater than her opponent's Step-1 performance. Improvement therefore cannot play any role.

Group) are respectively 0.15 ($p=0.17$) and 0.17 ($p=0.23$). Low-performing participants then do not adjust their decision to the level of the competition when the competition does not involve a subsequent performance. It therefore cannot be the case that low-performing subjects' adjustment of their tournament entry decision to the level of the competition is (only) driven by risk aversion, as no such effect is found in the decision to submit where risk aversion exactly plays the same role. Indeed, what seems to be driving our result is that low-performing participants from *Ability Group* see the possibility of improving their performance between Steps 3 and 4, while performance cannot be improved between Steps 3 prime and 4 prime (as for these steps only Step 1 performance matters). This can be seen as a taste for competition (subjects believe in their chances of improvement). This effect seems larger for men (although it does not reach significance).

For high-performing subjects, we only find a significant treatment effect for women. The diff-in-diff estimate for high-performing women is -0.41 ($p=0.01$) indicating that they adjust their decision to the level of the competition when considering whether to submit a past performance to a tournament. As such, the reference-group neglect (i.e. the under-adjustment to changes in the reference group one competes with (Camerer and Lovallo, 1999)) they exhibit in their decision to enter the tournaments may be due to an additional taste for competition after the receipt of positive feedback. High-performing women in the *Ability Group* treatment submit less to the tournament when they know that both themselves and their opponent are of high ability, while this is not the case for the entry choice (they actually enter a little more often in the second choice than in the first one). Women dismiss the information on the high level of the competition they will face when choosing whether to enter, but not when choosing whether to submit. For men, we find no significant treatment effect (the diff-in-diff estimate is -0.20 with $p=0.17$), as was the case for the decision to actually enter the tournament.

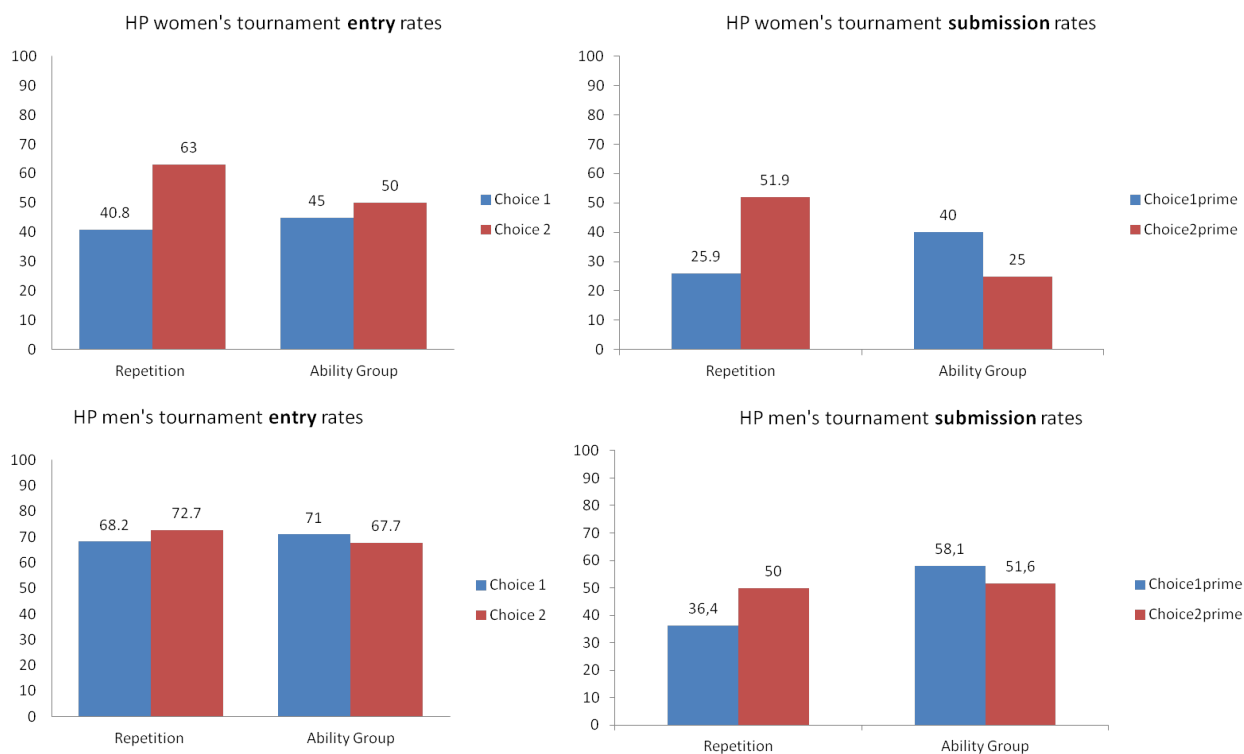


Figure 2.5. Proportion of high-performing (HP) women and men entering or submitting to the tournament (in %).

Figure 2.5 confronts the actual tournament entry rates with the submission rate, for high-performing participants (HP) by gender. We can see that while high-performing women are reluctant to submit to the tournament if they receive positive feedback in the *Ability Group*, this is not the case when they make their tournament entry decision. Positive feedback may then make them like to perform under the pressure of competition.

4.3. Regressions

Our experimental design allows us to determine how performance feedback, confidence and ability group affect the decision to enter a competition. We run linear-

probability regressions¹¹ (LPM) to evaluate these effects.

Our first concern is to see whether beliefs explain competitive behavior. We thus create the variable "beliefwin" (bw) as a proxy for the subject's beliefs concerning tournament success. We denote by b_{i_j} the beliefs elicited at round i of their performance being in the j th quartile. For Choice 1, beliefwin is equal to the belief of being above the median i.e. $bw = beliefsup = b_{1_4} + b_{1_3}$. If a subject thinks she has a 60% chance of being above the median she should also think she has a 60% chance of winning the tournament, as her opponent will be randomly chosen among the participants in her session. When having to decide for the second time whether to enter a tournament (Choice 2), the participants know whether they are above or below the median. The beliefs about the chances of winning will also depend on whether the subject was in the *Repetition* or in *Ability Group*. Participants in the *Ability Group* compete the second time against an opponent of the same ability group as their own, while, in *Repetition*, the opponent is again randomly chosen from all session participants. In *Repetition*, beliefwin will be equal to $beliefwin = 0.375 * b_{2_2} + 0.125 * b_{2_1}$ for low-performing subjects. We make the simplifying assumption that whenever a subject believes her performance belongs to a certain quartile, she actually thinks her performance lies exactly at the midpoint of this quartile. As such, a low-performing subject deciding whether to enter the second tournament should think she will beat all subjects from the worst quartile and half of the subjects in the second to last quartile (that is 37.5% of potential opponents) if she believes her performance belongs to the second to last quartile (which she thinks is b_{2_2} % likely). If she thinks her performance belongs to the worst quartile (which she thinks is b_{2_1} % likely), she would beat half of the subjects from the worst quartile (12.5%). Following the same reasoning, $beliefwin = 0.875 * b_{2_2} + 0.625 * b_{2_1}$ for high-performing subjects in *Repetition* and $beliefwin = 0.75 * b_{2_2} + 0.25 * b_{2_1}$ for both low- and high-performing subjects.

11. Probit regressions yield qualitatively similar results.

VARIABLES	Low-Performing Women			Low-Performing Men		
	(1)	(2)	(3)	(1)	(2)	(3)
Ability group	0.101 (0.113)	0.106 (0.105)	0.106 (0.105)	-0.033 (0.130)	0.059 (0.130)	0.067 (0.129)
Choice 2	-0.172* (0.100)	0.113 (0.149)	0.002 (0.094)	-0.029 (0.079)	0.119 (0.145)	0.224* (0.113)
Ability group*Choice 2	0.304** (0.148)	0.300** (0.144)	0.141 (0.143)	0.326** (0.131)	0.295** (0.132)	0.013 (0.168)
Beliefsup		0.007*** (0.002)			0.007*** (0.002)	
Beliefsup*Choice2		-0.005* (0.003)			-0.002 (0.002)	
Beliefwin			0.006*** (0.002)			0.008*** (0.002)
Constant	0.241*** (0.081)	-0.119 (0.098)	-0.109 (0.099)	0.588*** (0.086)	0.140 (0.184)	0.102 (0.182)
Observations	134	134	134	122	122	122
R-squared	0.102	0.170	0.170	0.060	0.157	0.145

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Ability group equals 1 for a treatment session, 0 for a Repetition session. Choice 2 is a dummy equal to 1 for the second choice to enter the tournament that is after feedback. Beliefsup is a continuous variable (between 0 to 100) corresponding to the stated belief of being above the median before receiving feedback.

Table 2.3: LPM for the tournament entry decision of low-performing women and men.

Table 2.3 shows the impact of beliefs on low-performing participants' tournament entry decisions. *Beliefsup* corresponds to the sum of the first-round elicited beliefs of belonging to the 4th and 3rd quartile, i.e. the beliefs, before receiving feedback, of being above the median. Ability group is a dummy for the participant being in an *Ability Group*, as opposed to a *Repetition*, session. Choice2 is a dummy variable which equals 1 if we consider the second decision to enter the step-4 tournament and 0 if we consider the first decision to enter the step-3 tournament. We have already discussed the results

of the regressions in column (1) in subsection 4.2 on the diff-in-diff estimators. The addition of Beliefsup and Beliefsup*Choice2 to the explanatory variables leaves the coefficient of Ability group*Choice2 basically unchanged for both low-performing men and women. The coefficient on Beliefsup*Choice2 is negative and significant at the 10% level for women, suggesting that the more confident low-performing women were to start with the more likely they are not to enter the tournament when having to decide for the second time.

Column (3) shows that when Beliefwin is added to the regressors, the coefficient on Ability group*Choice2 is lower and becomes insignificant for both genders. The adjustment of competitive entry to the level of their opponents is thus mainly driven by beliefs. Low-performing participants' lower confidence following negative feedback explains their diminished desire to enter the tournament afterwards.

We now consider the case of high-performing men and women (see table 2.4). In regression (2), the coefficient on Beliefsup*Choice2 is negative and significant at the 1% level for women but not for men. The more underconfident high-performing women were to begin with, the more likely they are to enter tournament the second time after receiving positive feedback. Together with the low-performing women in table 2.3, this suggests that women are prone to what we call a "surprise effect". Women are more likely to enter the tournament if they were initially pessimistic about their relative performance. This effect is not found for men.

Result 4: *Women react more strongly to feedback when they did not expect it. We call this the "surprise effect".*

The introduction of Beliefwin into the regressors in column (3) renders the coefficient on Ability group*Choice2 insignificant in both the male and female regressions.

Table 2.5 considers the impact of feedback on the tournament entry decision. The variable HighPerf is a dummy for the participant's step-2 performance being above the

VARIABLES	High-Performing Women			High-Performing Men		
	(1)	(2)	(3)	(1)	(2)	(3)
Ability group	0.043 (0.150)	0.054 (0.142)	0.052 (0.141)	0.028 (0.132)	0.019 (0.130)	0.017 (0.130)
Choice 2	0.222* (0.112)	0.880*** (0.201)	0.143 (0.102)	0.045 (0.104)	0.135 (0.244)	0.025 (0.107)
Ability group*Choice 2	-0.172 (0.143)	-0.187 (0.126)	-0.021 (0.139)	-0.078 (0.127)	-0.076 (0.127)	0.068 (0.138)
Beliefsup		0.008*** (0.002)			0.005 (0.003)	
Beliefsup*Choice2		-0.010*** (0.002)			-0.001 (0.002)	
Beliefwin			0.006** (0.002)			0.006* (0.003)
Constant	0.407*** (0.097)	-0.091 (0.140)	0.008 (0.148)	0.682*** (0.102)	0.327 (0.264)	0.247 (0.267)
Observations	94	94	94	106	106	106
R-squared	0.031	0.119	0.086	0.002	0.053	0.059

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.4: LPM for the tournament entry decision of high-performing women and men.

median of her session. This thus indicates the type of feedback (above the median vs. below the median) the participant received.

Women in the *Repetition* group react strongly to the type of feedback they receive, and enter significantly more following positive rather than negative feedback, as shown by the coefficient on Choice2*HighPerf in columns (1) and (2). This feedback reaction seems to work via beliefs, as the addition of Beliefwin to the regressors in column (3) knocks out the significance of Choice2*HighPerf. On the contrary, men do not seem to change their competitive behavior according to feedback.

In contrast to what we see in the *Repetition* group, men (but not women) change their competitive behavior according to their feedback in the *Ability Group* (see table

VARIABLES	<i>Repetition</i> women			<i>Repetition</i> men		
	(1)	(2)	(3)	(1)	(2)	(3)
HighPerf	0.166 (0.126)	0.097 (0.114)	0.098 (0.113)	0.094 (0.133)	0.031 (0.137)	0.024 (0.138)
Choice2	-0.172* (0.101)	0.321** (0.137)	0.061 (0.094)	-0.029 (0.079)	0.029 (0.156)	0.155 (0.118)
Choice2*HighPerf	0.395** (0.150)	0.466*** (0.148)	0.054 (0.138)	0.075 (0.130)	0.086 (0.142)	-0.129 (0.147)
Beliefsup		0.009*** (0.002)			0.005* (0.003)	
Beliefsup*Choice2		-0.009*** (0.003)			-0.001 (0.002)	
Beliefwin			0.009*** (0.002)			0.006** (0.003)
Constant	0.241*** (0.081)	-0.237** (0.099)	-0.229** (0.104)	0.588*** (0.086)	0.267 (0.191)	0.234 (0.197)
Observations	112	112	112	112	112	112

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Highperf is a dummy for subjects being above the median.

Table 2.5: LPM for the tournament entry decision among *Repetition* participants by gender.

2.6). The coefficient on Choice2*HighPerf is negative and significant in the male regressions (1) and (2), showing that men will choose to stay out of the competition the second time if they learned that they are above the median and their opponent will also be so. Again, this effect is mainly driven by beliefs. Women do not react to feedback when it also informs them about the level of their opponent. They are about as likely to choose tournament entry the second time if they learned that both they and their opponent are below the median as when they are both above.

The results from Tables 2.5 and 2.6 show that while low-performing men and women both adjust to the level of the competition while their high-performing counterparts do

VARIABLES	<i>Ability Group</i> women			<i>Ability Group</i> men		
	(1)	(2)	(3)	(1)	(2)	(3)
HighPerf	0.108 (0.138)	0.068 (0.138)	0.077 (0.138)	0.154 (0.128)	-0.040 (0.160)	0.077 (0.138)
Choice 2	0.132 (0.109)	0.437** (0.210)	0.139 (0.108)	0.296*** (0.105)	0.445*** (0.150)	0.139 (0.108)
Choice2*HighPerf	-0.082 (0.140)	-0.040 (0.134)	-0.040 (0.138)	-0.329** (0.128)	-0.248* (0.143)	-0.040 (0.138)
Beliefsup		0.005** (0.002)			0.007** (0.003)	
Beliefsup*Choice2		-0.006* (0.003)			-0.003 (0.002)	
Beliefwin			0.004** (0.002)			0.004** (0.002)
Constant	0.342*** (0.079)	0.047 (0.128)	0.110 (0.122)	0.556*** (0.098)	0.198 (0.180)	0.110 (0.122)
Observations	116	116	116	116	116	116
R-squared	0.017	0.053	0.043	0.049	0.134	0.043

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.6: LPM for the tournament entry decision among *Ability Group* participants by gender.

not, the reasons for doing so are different. While women mainly react to the feedback on their own performance, men focus more on the information regarding the level of their opponent.

4.4. Welfare analysis

We now turn to the consequences of competition behavior on welfare. More precisely, we are interested in whether choices maximize expected payoffs.

To calculate the expected payoffs from entering the standard tournament (i.e. the tournament where the opponent is randomly drawn among all other participants in

the session), 100,000 performances were drawn by sampling with replacement from the Step-2 performances of our 228 participants. For each level of performance, the probability of tournament success was computed by calculating the number of times out of 100,000 this given performance exceeded the opponent's performance. Similarly, for ability-grouping tournament success for low-performing participants, 100,000 performances were drawn from the Step-2 performances of the potential opponents, i.e., participants whose Step-2 performance was also below the median. We then calculate, for each performance level, the number of times out of 100,000 this given performance exceeded the opponent's performance. The same method is used to compute ability-grouping tournament success for high-performing participants. We then compare, for each performance level, the payoff from choosing the piece rate to that from entering the tournament. This tells us which participants would have maximized their payoffs by entering the tournament, which we compare to participants' actual decisions. Given the distribution of Step-2 performances, all participants with a performance of 9 or more have a higher expected payoff from standard tournament entry than from the piece rate. For instance, if participants expect their Step-3 performance to be the same as in the Step-2 tournament, 53.5% of participants should enter the Choice-1 tournament. However, taking into account actual Step-3 performances, which are slightly better than Step-2 performances, 61% of participants would have gained from choosing the tournament. In the same way, all low-performing (high-performing) participants with a performance at least equal to 7 (11) should enter the Step-4 ability group tournament in the *Ability Group* treatment. Figures 2.6 and 2.7 show the percentage of participants by gender, treatment and ability level who enter each type of tournament as well as the percentage who would have profited from doing so, both if their performance remained at the Step-2 level and at their actual step-3 performance.

From a welfare point of view, we are most interested in comparing the observed rate of entry to that predicted by participants' actual performances.

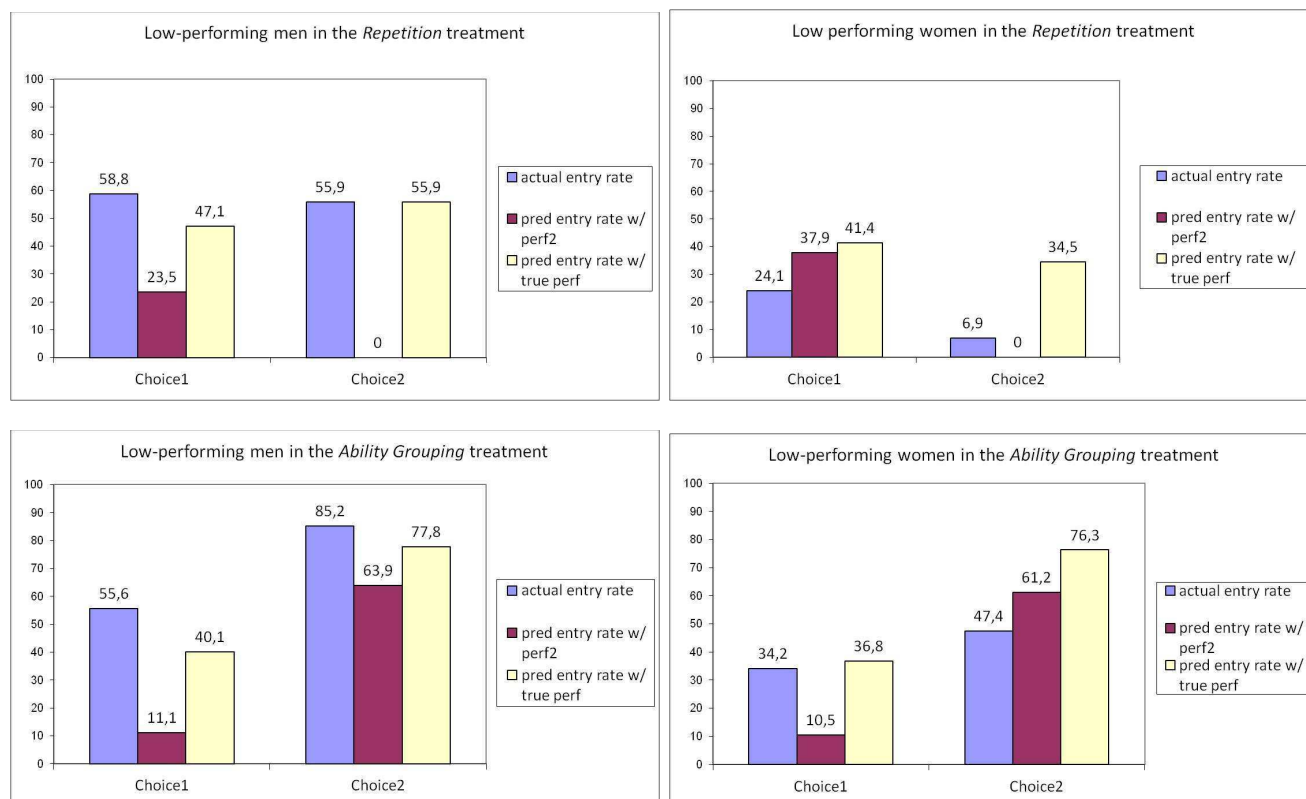


Figure 2.6: Actual and predicted entry rates of low-performing participants by gender and treatment

Figure 2.6 shows that while low-performing men enter in about the payoff-maximizing proportion, not enough low-performing women choose the tournament the second time (the two-tailed z-test of the proportion actually entering to that predicted by true performances yields $p=0.01$). In *Repetition*, women react to negative feedback as if they expect no performance improvement between Steps 2 and 4, while their performance does actually improve. In *Ability Group*, they do not sufficiently incorporate that their opponent will also be of low ability.

Among high-performing participants, men tend to not enter as much as predicted when making their first choice ($p=0.06$ and 0.05 respectively in the *Repetition* and *Ability Group*). In *Repetition*, high-performing women do not enter enough both before

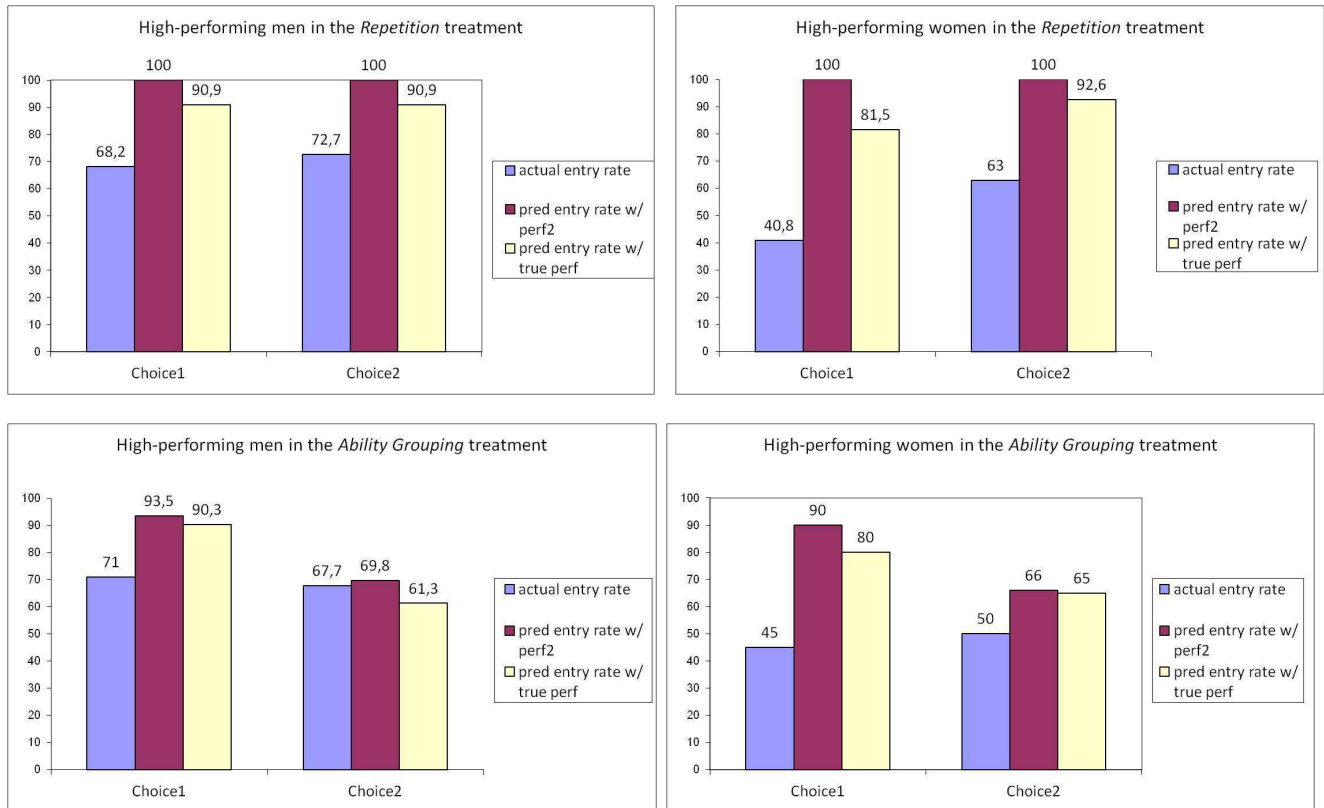


Figure 2.7: Actual and predicted entry rates of high-performing participants by gender and treatment

and after receiving positive feedback ($p < 0.01$ in both cases). In *Ability Group*, high-performing women do not enter enough when making their first choice but do not enter significantly less than predicted the second time. Feedback seems to improve men's decisions in terms of maximizing expected payoffs, contrary to women who do not enter tournament enough even after receiving additional information.

Result 5: *Most of the time, men enter tournaments in about the proportion which maximizes their payoffs. Women tend not to choose the tournament enough and when the information concerns only their own performance, they do not enter more following positive feedback but enter even less after negative feedback.*

This result is in line with [Datta Gupta, Poulsen and Villeval \(2012\)](#).

5. Discussion and conclusion

This paper has shown that subjects update their beliefs following performance feedback more strongly than would a Bayesian agent. Both men and women are more pessimistic than a Bayesian agent following negative feedback; we find the opposite effect after positive feedback. Both effects are stronger for women than for men. We also show that low-performing participants adapt their tournament entry decision to the level of the competition, while high-performing participants do not.

Concerning feedback reaction in terms of competitive entry, men and women do not react in the same way. While women are especially sensitive to information on their own performance level, men react more strongly to the level of their competitors. Feedback does not therefore seem to be processed in the same way by men and women. As a result, women react very strongly to feedback received on relative performance, while men respond more to information on the performance they will have to beat, which is fixed by design. One important point is that the effects of feedback and information concerning one's opponent's level on tournament entry decisions mostly transit through the subjective belief of winning the tournament. We furthermore find what we call a "surprise effect" for women: they react more strongly to feedback when they did not expect it.

Our welfare analysis shows whether participants maximize their payoffs by their tournament entry decisions. Did they lose money by making the wrong choice? We show that while men enter most of the time in about the payoff-maximizing proportion, not enough women choose the tournament. This can be explained by women giving too much weight to the negative feedback they receive and not taking into account the fact that their performance can (and does) improve over time. In *Repetition*, women react to the negative feedback as if they expect no progress in their performance between Steps 3 and 4, while their performance does actually improve. In the *Ability Group*, low-

performing women do not sufficiently take into account the fact that their opponent will also be of low ability. Only internal and self information seems to matter in their decision. Implementing an ability group tournament seems to overall bring the proportion of entrants closer to the optimal proportion given true performances. This effect is however counter-balanced by the provision of a negative feedback to women.

The main message of this paper is that not only beliefs but also the way an agent updates his beliefs when receiving information about relative performance play a major role in the tournament entry decision. People, and especially women, overreact to the information they receive. Men seem to internalize more the information on the level of the competition they will face, and take this into account in their decision process. Women however overreact to feedback, and even when they know the level of their opponent, they put too much weight on their feedback. A policy recommendation could be to refrain from sending precise feedback about relative performance and focus rather on the possibility for improvement. This could for instance be implemented in firms by setting goals for employees rather than focusing on bad relative performance during annual interviews. Indeed, the receipt of negative evaluations can lead to a decrease in motivation and discourage the provision of effort. This may especially be important for schoolgirls and boys, for whom the way beliefs are formed and processed may not yet be set in stone. The long-term effect of ability group on performance through peer effects must also be considered.

B. Appendix

B.1. Tables

Variable	Modality	
Age		25.8
Discipline	Economics	31.1%
	Science	3.1%
	Mathematics	2.2%
	Others	63.6%
Study level	Bac or less	3.95%
	Bac+1 to Bac+2	40.8%
	Bac+3	18.86%
	Bac+4 to Bac+5	34.21%
	More	2.19%
Father's education	Bac or less	35.53%
	Bac+1 to Bac+3	28.95%
	Bac+4 and more	35.52%
Mother's education	Bac or less	39.04%
	Bac+1 to Bac+3	34.21%
	Bac+4 and more	26.75%
Already participated in an experiment	Yes	73.25%

Table 2.7: Descriptive statistics.

The Baccalauréat or "Bac" is an academic exam (and diploma) that French pupils take at the end of high school. Passing the Bac allows students to continue to higher education. Therefore, for instance, Bac+1 refers to a level of education of 1 year following the acquisition of the Baccalauréat.

B.2. Instructions

The experiment consists of six steps. Before each step, we will carefully explain to you what the step is about and you will have the opportunity to ask as many questions as you require¹². Please remember that you are not allowed to communicate in any way with one another. At the end of the experiment one of the six steps you will have completed will be randomly chosen to determine your payoffs.

Step 1. Piece Rate: In step 1, you will have five minutes to solve as many sums of five two-digit numbers as you can. You are allowed to use the scratch paper you have been given. If step 1 is the randomly-chosen step for payment, you will receive €0.50 per correctly-solved sum. At the end of step 1, a screen will indicate how many sums you solved correctly.

NEXT SLIDE

Step 2. Individual Tournament: You will have five minutes to solve as many sums of five two-digit numbers as you can. If step 2 is chosen for payment, you will receive €1 per correct answer if you solved more sums than a randomly-chosen opponent present in the room, otherwise you will receive nothing. You will earn €0.50 per sum correctly solved in case of a tie.

At the end of step 2, a screen will indicate how many sums you solved correctly but you will only find out whether you won this tournament at the end of the experiment.

NEXT SLIDE

Compare your performance to that of the others: There are four groups of performance in the session (Q1, Q2, Q3, Q4), each consisting of 25% of the participants. The first group Q1 contains the 25% best participants, the second group Q2 the following 25% etc.

12. Subjects were not given the instructions. We use a Powerpoint presentation to explain each step before it begins.

You have to state your confidence about belonging to each of the four performance groups based on your step-2 performance. You should give a probability in % about your belonging to each group. Hence, you have to supply four probabilities, the sum of these four being equal to 100%.

Payoff rule: We use an incentive rule so that you tell us what you really think. Take the second best group Q2: you think you have an $x\%$ chance of belonging to this second group and inform us of this probability. The computer draws a random number y that belongs to the interval $[0; 100]$.

- If $y < x$, you earn €1 if your performance exactly belongs to the second group Q2.
- If $y > x$, you earn €1 with $y\%$ probability.

The same rules applies to the three other groups.

Example:

You answered you think you have a 30% chance that your step-2 performance belongs to Q2.

If the randomly-drawn number is smaller than 30, let's say 19, then you win €1 if your step performance belongs to Q2.

If the randomly-drawn number is greater than 30, let's say 54, then you win €1 with a 54% chance.

NEXT SLIDE

Step 3. Choice between Piece Rate and Individual Tournament: Before performing your five minutes of sums, you will have to choose whether you want to be paid according to the Piece Rate (€0.50 per correct answer) or the Individual Tournament compensation scheme.

If you choose the Piece Rate, you will receive €0.50 per sum correctly solved during step 3. If you select the Tournament, you will receive €1 per correct answer if your step-3

performance exceeds the step-2 performance of a randomly-chosen opponent, otherwise you will receive nothing. You will earn €0.50 per sum correctly solved during step 3 in case of a tie.

At the end of step 3, a screen will indicate how many sums you solved correctly but you will only find out whether you won this tournament, if you chose this option, at the end of the experiment.

NEXT SLIDE

Step 4¹³. Choice between *submitting* your step-1 performance to a Piece Rate or an Individual Tournament: There are no sums to be solved here, as the performance which will determine your payoffs is your step-1 performance.

If you choose to submit your step-1 performance to the Piece Rate, you will receive €0.50 times your step-1 performance. If you choose to submit your step-1 performance to the individual tournament, you will receive €1 per sum correctly solved in step 1 if you solved more sums in step 1 than your randomly-chosen opponent, otherwise you will receive nothing.

You will earn €0.50 per sum correctly solved during step 1 in case of a tie. You will only find out whether you won your tournament, if you choose to submit your Step-1 performance to the tournament, at the end of the experiment.

NEXT SLIDE

Feedback: The computer tells you if your **step-2 performance** was situated above or below the median performance in the session.

- If your step 2 score was above the median (top half), you belong either to Q1 or Q2.
- If your step 2 score was below the median (bottom half), you belong either to Q3 or Q4.

13. The step which was presented as step 4 to the participants is labelled "step 3 prime" in the article text.

Compare your performance to that of the others (round 2): You now have to state your probabilities (in %) of belonging to each of the two possible groups (Q1 and Q2, or Q3 and Q4) according to the feedback you received. The sum of these two percentages must equal 100. We use the same payoff rule as previously.

NEXT SLIDE

Step 5¹⁴. Choice between a Piece Rate and a Tournament:

- In the *Repetition Sessions*: The step follows exactly the same procedure as in step 3. Before performing your five minutes of sums, you will have to choose whether you want to be paid according to the Piece Rate (€0.50 per correct answer) or the Individual Tournament compensation scheme.

If you choose the Piece Rate, you will receive €0.50 per sum correctly solved during step 3. If you select the Tournament, you will receive €1 per correct answer if your Step-3 performance exceeds the step-2 performance of a randomly-chosen opponent, otherwise you will receive nothing. You will earn €0.50 per correctly-solved sum during step 3 in case of a tie.

- In the *ability group Sessions*: Before performing your five minutes of sums, you will have to choose whether you want to be paid according to the Piece Rate (€0.50 per correct answer) or the Individual Tournament compensation scheme. If you choose the Piece Rate, you will receive €0.50 per correctly-solved sum during step 3. If you select the Tournament, you will be randomly paired with an opponent who is **in the same ability group** as you. You will receive €1 per correct answer if your step-3 performance exceeds the step-2 performance of this same ability randomly-chosen opponent, otherwise you will receive nothing. You will earn €0.50 per correctly-solved sum during step 3 in case of a tie.

At the end of step 5, a screen will indicate how many sums you solved correctly, but

14. The step which was presented as step 5 to the participants is labeled "step 4" in the article text.

you will only find out whether you won your tournament, if chosen, at the end of the experiment.

NEXT SLIDE

Step 6¹⁵. Choice between *submitting* step 1 performance to a Piece Rate or an Individual Tournament: There are no sums to be solved here, as the performance which will determine your payoff is your step-1 performance.

- In the *Repetition* Sessions: If you choose to submit your step-1 performance to the Piece Rate, you will receive €0.50 times your step-1 performance.
If you choose to submit your step-1 performance to the individual tournament, you will receive €1 per correctly-solved sum in step 1 if you solved more sums in step 1 than your randomly-chosen opponent, otherwise you will receive nothing. You will earn €0.50 per correctly-solved sum during step 1 in the case of a tie.
- In the *ability group* Sessions: If you choose to submit your step-1 performance to the Piece Rate, you will receive €0.50 times your step-1 performance.
If you choose to submit your step-1 performance to the individual tournament, you will receive €1 per correctly-solved sum in step 1, if you solved more sums in step 1 than your randomly-chosen opponent who belongs to the same ability group, otherwise you will receive nothing. You will earn €0.50 per correctly-solved sum during step 1 in the case of a tie.

You will only find out whether you won this tournament, if you choose to submit your step-1 performance to the tournament, at the end of the experiment.

15. The step which was presented as step 6 to the participants is labelled "step 4 prime" in the article text.

Part II

Creativity, Personality and Teenagers' Behaviors

CHAPTER 3

CREATIVITY, PERSONALITY AND SCHOOLING
ACHIEVEMENT: AN EXPLORATORY STUDY ON
FIELD DATA

This chapter is a joint work with Jean-Louis Tavani and Maud Besançon¹.

1. It has been submitted to Education Economics.

1. Introduction

Education has a major impact on economic outcomes, especially on the labor market. Schooling choices should lead individuals to optimal choices with respect to their abilities and environment. Differentials in this achievement are driven by many factors such as individual causes (own abilities, social background) as well as environmental causes (peer effects, neighborhood effects) or causes related to the schools inner workings (public vs. private, credit constraints, early or late specialization, available information etc...).

For a long time, cognitive abilities were assumed to be the only reason of success in school and in the labor market². However there are non-cognitive attributes, that are academically and occupationally relevant skills and personality traits, which are not correlated with intellectual skills (such as IQ). They include a range of personality and motivational habits and attitudes that facilitate performing well in school. Perseverance, motivation, self-control, and other aspects of conscientiousness are often found to be the most significant (Borghans et al., 2008). In the recent economic literature, it has been shown that there is indeed an effect of cognitive and non-cognitive abilities on educational outcomes and wages (Bowles and Gintis, 1975, Bowles, Gintis and Osborne, 2001, Heckman, 2006, Heckman, Stixrud and Urzua, 2006). Bandura (1993) sheds light on how perceived self-efficacy would affect cognitive and motivational processes, showing then that non-cognitive abilities play a major role in explaining personal achievement. Non-cognitive skills are also good predictors for some bad behaviors such as drug addiction or early pregnancy, whereas cognitive skills are not (Heckman, 2006).

This paper presents an exploratory study and aims to go deeper into the analysis of explaining schooling achievement. It investigates the influence of creative potential, personality and other individual characteristics on academic achievement of middle

2. We indeed hypothesize here that the correlation between wages and schooling grades exists, through higher education attainment, and is positive (Weiss, 1995, Altonji, 1995).

school students. It mainly focuses on the individuals' abilities.

Cattell (1979) says that personality "permits a prediction of what a person will do in a given situation". The Five-Factor Model (FFM) of personality traits (Big Five) is certainly the dominant theory in personality research (McCrae and Costa, 1987, John and Srivastava, 1999). The personality vocabulary provides a finite set of attributes for personality description (Goldberg, 1981). The Big Five model was discovered in analyses of several languages (Saucier and Goldberg, 1998) and a significant amount of research has confirmed its great utility and its accuracy. This model proposes a structure of human personality in five dual dimensions: agreeableness, conscientiousness, extraversion (vs. introversion), emotional stability (vs. neuroticism), and openness. Psychological literature has investigated the influence of personality on academic achievement³ finding that conscientiousness is the main dimension that positively predicts academic achievement. To a larger extent, some personality traits seem to have a less consistent effect. It is the case of agreeableness (Trapmann et al., 2007), openness to new experiences and extraversion (O'Connor and Paunonen, 2007). Although extraversion seems to be negatively correlated with higher education, it is positively correlated with primary and secondary schooling (*ibid.*). As our subjects are teenagers, we use the Big Five inventory targeted for teenagers called *The Brief Big Five* (BB5) (Barbot, 2012). This questionnaire allows us to obtain the scores for these same five dimensions.

While many studies have been conducted to investigate the effect of personality on academic achievement, studies on the role of creativity are more rare. Creativity can be both considered as a cognitive or non-cognitive ability that requires knowledge, intellectual style, personality, motivation and environmental context (Sternberg and Lubart, 1995). It is also defined as the capacity to produce something new, original and adapted to the constraints of a given situation (Amabile, 1996, Runco, 2004,

3. See O'Connor and Paunonen (2007), Laidra, Pullmann and Allik (2007), Komarraju et al. (2011), Poropat (2009), Richardson, Abraham and Bond (2012), Trapmann et al. (2007)

Lubart, 1994). The differences observed between individuals result from a combination of cognitive, conative⁴ and environmental factors (Lubart et al., 2003). The influence of environment occurs in different spheres: the family, the school or work environment and the cultural context in which the person evolves. School environment and teaching methods can contribute to the development of creativity which can lead to greater innovations (Besançon and Lubart, 2008). This new dimension is rarely studied in economics although it strongly predicts innovation as this latter represents the successful implementation of creative ideas within an organization (Majaro, 1992, Antonites and Van Vuuren, 2005). Because creativity and innovation are related, and because innovation is a keystone of performing firms, the role of creativity as early as during childhood becomes an important dimension. The underlying idea is that non-creative pupils/students could become non-creative workers which would thus limit future innovations in the firms. To our knowledge, there are only two papers that are interested in the effect of creativity and personality on schooling achievement: Furnham, Zhang and Chamorro-Premuzic (2006), Sen and Hagtvet (1993) who both find no significant relationship between creativity and intelligence (measured with psychometric and IQ tests). Therefore, the original feature of our paper is that we also add creativity scores in our estimations to see its capacity to explain school grades.

As this paper analyzes how much creativity can be associated to grades, it gives a glance at the role of creativity in the French educational system. The role of creativity in education has become these last ten years a quite important issue. Reports from the French Institute of Education (Ifé) question and develop the question of creativity in French schools. In their report of January 2012⁵ they describe the French educational system as formalized and organized in order to relay existing knowledge but that is usually suspected to be rewarded on conformity. This goes against a creativity learning.

4. As opposed to cognitive factors, conative factors refer to personality traits and to motivation.

5. This report can be found in French at <http://ife.ens-lyon.fr/vst/DA-Veille/70-janvier-2012.pdf>.

However, especially in the actual society with a real turnover on the labor market, skills as risk-taking, flexibility and creativity are very valuable. Abilities required on the labor market are more and more linked to creativity such as problems resolutions, abilities to analyze or self-management. The Ifé proposes a trade-off between innovation in terms of pedagogical tools based on creativity learning and traditional evaluation. This idea is still questioned in France, especially because schooling programs are totally State established. This implies that changing any aspects of schooling pedagogy for most of the public schools is an administrative and legislative process (private schools or schools with specific pedagogies are excluded).

In order to identify a relationship between creativity and schooling achievement in addition to traditional non-cognitive measures (personality traits) and individual characteristics, we collected data in the field. We went into classrooms of a middle school located in a Parisian suburb. Our sample consists of six 9th grades classes that took the BB5 and a measure of risk aversion with lotteries (Holt and Laury, 2002), and we measured the pupils' creativity.

We do not consider creativity as a unique and general variable. Creativity is a complex concept and each individual expresses her creativity under different domains of expression and factors. We hence use a multivariate approach (EPoC, Lubart, Besancon and Barbot (2011)): someone could have a high potential in one domain but not in another one (writing vs. graphic for instance). Hence, measures of creativity are separated into four scores of two types of creativity: divergent thinking and integrative thinking, for both graphic and verbal skills. Integrative thinking is the activity of combining elements in new ways. It is an associative thinking, selective comparison and combination of ideas that allows synthesis of various heterogeneous elements to converge into a unique and original production. Divergent thinking is the process of expanding the range of solutions in creative problem solving that includes factors such

as flexibility and selective encoding⁶.

To measure school achievement we obtained the pupils' report cards from the school in order to have the average grades for every subject. We estimate the grades of the mandatory courses that are math, physics and chemistry, biology, French, world languages (WL), history and geography, music, art, IT and physical education (PE). Math, French and world languages can be considered to be the fundamental courses by the educational system, followed by physics and chemistry, biology, history and geography. The four others can be categorized as secondary. We also obtained the success rate of the final middle school exam (Brevet des Collèges) and we are thus able to estimate its probability of success.

Based on our specific sample, we find a weak correlation between our creativity measures and the personality traits but with openness, which reinforces the independent role of creativity and also the fact that a higher openness to experience is generally associated with a higher creative potential (McCrae and Costa, 1987).

The main result of our paper is that creativity has an ambiguous effect on school achievement. First, verbal divergent thinking negatively predicts almost all of the subjects' grades. It is the only creativity subtype that has a significant effect on grades. These result sheds light on the fact that in order to obtain good grades at school it is necessary to restrain divergent thinking. Second, when looking at the probability to pass the final middle school exam (brevet des collèges), we find that verbal divergent thinking has no more impact, but that verbal integrative thinking predicts positively this probability.

Openness, and to a lesser extent conscientiousness, are the only personality traits that exhibit significant positive estimates on grades, but no effect of personality is found on the probability of succeeding the exam. An inconsistency variable, coming from our risk aversion measures is negatively associated with scientific grades.

6. See Sternberg and Kaufman (2010), Feldhusen and Goh (1995) for a review of the different creativity definitions as well as different creativity assessment measures.

Overall, girls have significantly higher grades than boys, but in sports. This might explain the fact that contemporary girls often choose different paths and stay in schools longer than boys⁷. Moreover, girls have higher scores of verbal thinkings (both divergent and integrative) compared to boys. An interesting results emerges from the analysis on the exam's probability of success: all things being equal, boys have a higher probability of passing the exam, although we show that girls have higher grades during the school year. We discuss different explanations for this result.

This paper is organized as follows: section 2 describes our data and how we collected it. Section 3 describes our results. Finally section 4 discusses the results and section 5 concludes.

2. The data

Data was collected on six 9th grade classes from a middle school in Rosny-sous-Bois (a Parisian suburb) in January and February 2012. The whole procedure has four steps. The first two steps correspond to the measurement of creativity, the third step corresponds to the personality test (the BB5 based on the Big Five inventory) and a measure of risk aversion based on the Holt and Laury (2002) task. The last one consists in gathering the schooling grades of each student in each subject that reflects schooling achievement and their cognitive skills. The following subsections describe each step.

2.1. Measuring creativity using the Evaluation of Potential Creativity (EPoC)

There exists many different test of creativity, and it has proven to be a difficult psychological concept to measure. Some tests are based on completing tasks (the Torrance Test of Creative Thinking for instance, Torrance (1968)) and some are self-assessed

⁷. We are not taking into consideration here the stereotype threat girls and boys may encounter when choosing specialities.

(The Adjective Check List, Zuckerman and Lubin (1965)). When based on performing tasks, the existing tests yield scores of divergent thinking (verbal and graphic), but not on integrative thinking. To our opinion, it is important to consider both dimension as they represent two types of skills for one's creative potential equally important.

EPoC (Lubart, Besancon and Barbot, 2011) is a procedure to measure the creative potential of pupils from elementary to middle school. It is a synthesis and extension of several traditions of measurement, which is based on a current theoretical framework envisioning creativity as multi-faceted, domain specific construct that involves many components. Through this procedure, it is possible to categorize the numerous micro-processes involved in creative potential into two main sets, called divergent-exploratory processes, and convergent-integrative processes. Moreover, as creativity is domain-specific, it is important that measures of creative potential take into account the domain of creative expression. Consequently, EPoC measures both sets of micro-processes: the divergent thinking (DT) and the integrative thinking (IT), based on two different domains of expression: verbal and graphic. This is contrary to other existing measurement tools that only focus on a single component of creativity with a global approach and tend to generalize the observed results to any domain of creative expression. A detailed description of the tasks and some pupils' outputs can be found in the appendix of part 2 of this dissertation.

Divergent thinking (DT) is a thinking process consisting in generating a maximum of creative solutions. The EPoC procedure embodies two types of DT:

- Graphical Divergent Thinking (hereafter *GDT*): pupils are asked to produce a maximum of original drawings in 10 minutes based on a simple shape.
- Verbal Divergent Thinking (hereafter *VDT*): the experimentalist gives a beginning or an end to a story and pupils are asked to produce a maximum number of story endings or beginnings in 10 minutes.

Integrative thinking (IT) is a cognitive activity that consists of combining many ele-

ments. Two types of IT are distinguished:

- Graphical Integrative Thinking (hereafter *GIT*): From ten drawings, the pupils are asked to produce one unique drawing, the most original, using at least three drawings from the list in 10 minutes.
- Verbal Integrative Thinking (hereafter *VIT*): The experimentalist gives a story title and the pupils have 10 minutes to finish the story in the most original manner.

The external validity of the procedure has been confirmed by [Lubart, Besancon and Barbot \(2011\)](#) by measuring creativity with EPoC as well as replicating a measure of creativity by [Torrance \(1962\)](#): this test is also known as the "cardboard box" where subjects have 10 minutes to propose a maximum of possible utilisation of the box, hence categorized as a divergent thinking task. It showed a high and significant correlation with the divergent thinking tasks of EPoC. There are well known results between IQ tests and creativity measures showing a low correlation between these measures. Hence, the IQ test were run on children who also took the EPoC tests and confirm this result. Finally, subjects also took the BB5 test (described in the next session). [Lubart, Besancon and Barbot \(2011\)](#) found a correlation between the trait of openness and creativity dimensions which is consistent with the results from [McCrae and Costa \(1987\)](#) saying that this personality trait facilitate the efficient use of divergent thinking so it can be as creative as possible. The external validity of the EPoC procedure is thus well established.

Two sessions have been organized to measure creativity⁸. During each session, each type of creative thinking was measured (see figure 3.5, 3.6, 3.7 in the appendix of this chapter for different outputs of VIT, VDT and GIT). Each session lasted on average 50

8. A different version (A and B) is used in each session. This is aimed to obtain robust scores of creativity. The tasks are globally the same but the content differs (type of drawings, titles of the stories...).

minutes and were run in the classroom with paper, pencils and felt pens. To summarize, in the DT tasks, pupils were asked to produce a maximum of ideas, responding to a unique stimulus. In the IT tasks, they had to produce a more elaborated synthetic solution.

Creativity scores for each type are based on the number of outputs (for DT tasks) and the degree of originality. The higher the score, the higher the individual's creative potential. In order to simplify the analysis, we use standardized scores of these four measures. 81 pupils fulfilled the whole test (on both weeks).

Table 3.1 report the correlation coefficients between the four score. They are all positively, significantly but moderately correlated with a coefficient varying from 0.18 to 0.4.

	GIT	VIT	GDT	VDT
GIT	1.0000			
VIT	0.3859 (0.00)	1.0000		
GDT	0.2509 (0.02)	0.2596 (0.02)	1.0000	
VDT	0.1899 (0.09)	0.1990 (0.07)	0.2631 (0.02)	1.0000

Table 3.1: Correlations table between the four scores (p-values in parentheses).

The distribution of the four creativity scores can also tell how much creative our subjects are. Figure 3.1 shows the distribution of the four creativity scores. We can see that higher integrative thinking is more frequent compared to divergent thinking whose distribution is more shifted to the left on the distribution graph. It already signifies

that children are better at integrative thinking than at divergent thinking. [Lubart, Besancon and Barbot \(2011\)](#) suggest that integrative thinking (especially verbal) are developed earlier than divergent thinking. Even though it is still not totally confirmed, this might explain why we observe this kind of distribution. The remaining question is to see whether these creativity dimensions have an influence on schooling achievement.

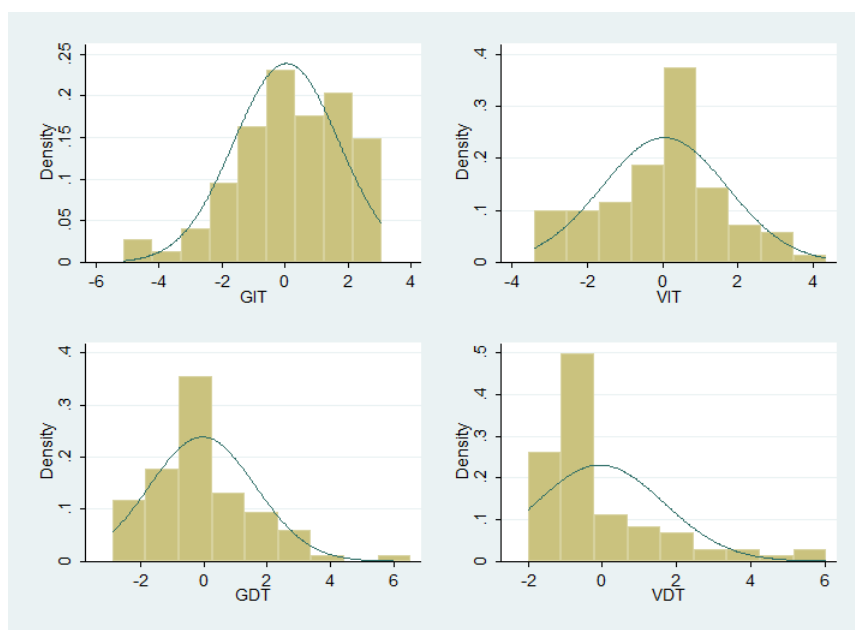


Figure 3.1: Distribution of the four standardized creativity scores

2.2. The BB5 questionnaire

In this study, as to measure the five personality traits: Agreeableness, Conscientiousness, Extraversion, Emotional Stability and Openness (as in the Big Five inventory) we use the Brief Big Five measure ([Barbot, 2012](#)) which has been specifically built for children and teenagers. This questionnaire consists of a list of 100 adjectives. For each adjective, the pupil declares on a five-point Likert scale if the adjective describes her/him totally or not at all, with three intermediate possibilities. Pupils were allowed to ask questions if they did not understand an adjective. We obtain a unique score for

each trait. It allows highlighting the personality traits that characterize the individual.

After the BB5 questionnaire, we asked pupils traditional demographic questions such as gender, age, parental education etc... 99 pupils took this test on a computer notebook, in the classroom.

There are validity criteria based on missing entries, "non-positioning/doubts" (this is the tendency to give central answers rather than positioning oneself on agreement or disagreement), and the tendency of agreement (a high frequency of total agreement or disagreement). These last two criteria yield a score allowing to determine valid scores. Three pupils did not finish the questionnaire and five did not have valid final scores. Consequently, 91 observations are usable for analysis based on the BB5 questionnaire. Table 3.8 in the appendix shows the personality trait scores by gender. We find that girls are significantly more open and agreeable than boys, and slightly more conscientious.

2.3. Inconsistency based on a risk aversion measure

We implemented a risk aversion elicitation rule as in Holt and Laury (2002) right after the BB5 questionnaire. They had ten choices to make between two lotteries (see figure 3.2). Probabilities were the same for each choice, only the amount of money changed.

Lotteries were presented as scratching cards where the probability of winning were presented as chances of winning. Pupils were allowed to ask questions about the task which allows us to check that overall they understood quite well the instructions and the task. We are interested here in the inconsistency yielded by this measure. The rate of inconsistent choices we observe is 26.3%. Inconsistent subjects are those who exhibit multiple switches or inconsistent choices (like choosing option A in the last row)⁹. This inconsistency rate does not significantly differs from previous studies made

9. We allow subjects who switch three times to be consistent considering that they might be indifferent between the first and the third swith. For these individuals, their certainty equivalent is situated



Figure 3.2: Lottery choices

on adults (see Ballinger and Wilcox (1997), Blavatsky (2010), Lévy-Garboua et al. (2012), Loomes and Sugden (1998)). In this study, we use the inconsistency as a proxy of a certain type of cognitive ability relying on the assumption that subjects who are considered as inconsistent either did not understand the task or have troubles with probabilities computation.

2.4. Grades collected

In order to measure schooling achievement, we collected the average grades of every pupil for each subject and each trimester of their 9th grade. We use the mandatory courses of middle school that are: mathematics, physics and chemistry, biology, French, history and geography, world languages, arts, music, IT and physical education (PE). We distinguish these courses into three categories: the scientific subjects, the humani-

between the first and the third row.

ties and the secondary ones (see table 3.2).

Category	Course	N
Sciences	Mathematics	97
	Physics and Chemistry	95
	Biology	92
Humanities	French	97
	World language (WL)	96
	History and geography	97
Secondary	Arts	69
	Music	83
	IT	96
	Physic Education (PE)	90

Table 3.2: Mandatory courses a middle school pupil attends and number of observations for each subject.

As table 3.2 shows, we were able to gather between 92 and 97 observations for the fundamental courses. However, even though the secondary subjects are mandatory, it is possible that some pupils skip them explaining the low number of observations for these courses. This may indicate that these courses are underestimated and considered as less important than the others.

There are six 9th grade classes named A to F. In France, pupils are graded on a 20 points scale. Figure 3.3 shows the distributions of the standardized grades in each subject which are close to normal (but for math, physics and chemistry, and French). In addition, table 3.9 in the annex reports all the average grades in each class for each subject considered. There are some classes having higher average grades than in the others. This can be due to the endogeneity of notations (peer effects¹⁰, teacher effects). We check for each teacher in each class and subject and find that when there are differences between average grade classes, it is not because one teacher grades his classes differently but because the teacher is different. Hence, when classes have the

10. Hoxby (2000) and Hanushek et al. (2003) finds that peer achievement has a positive effect on students' own achievement.

same teacher, the average grades are close to be the same. In order to overcome any peer effects, class fixed effects will be included in the estimations.

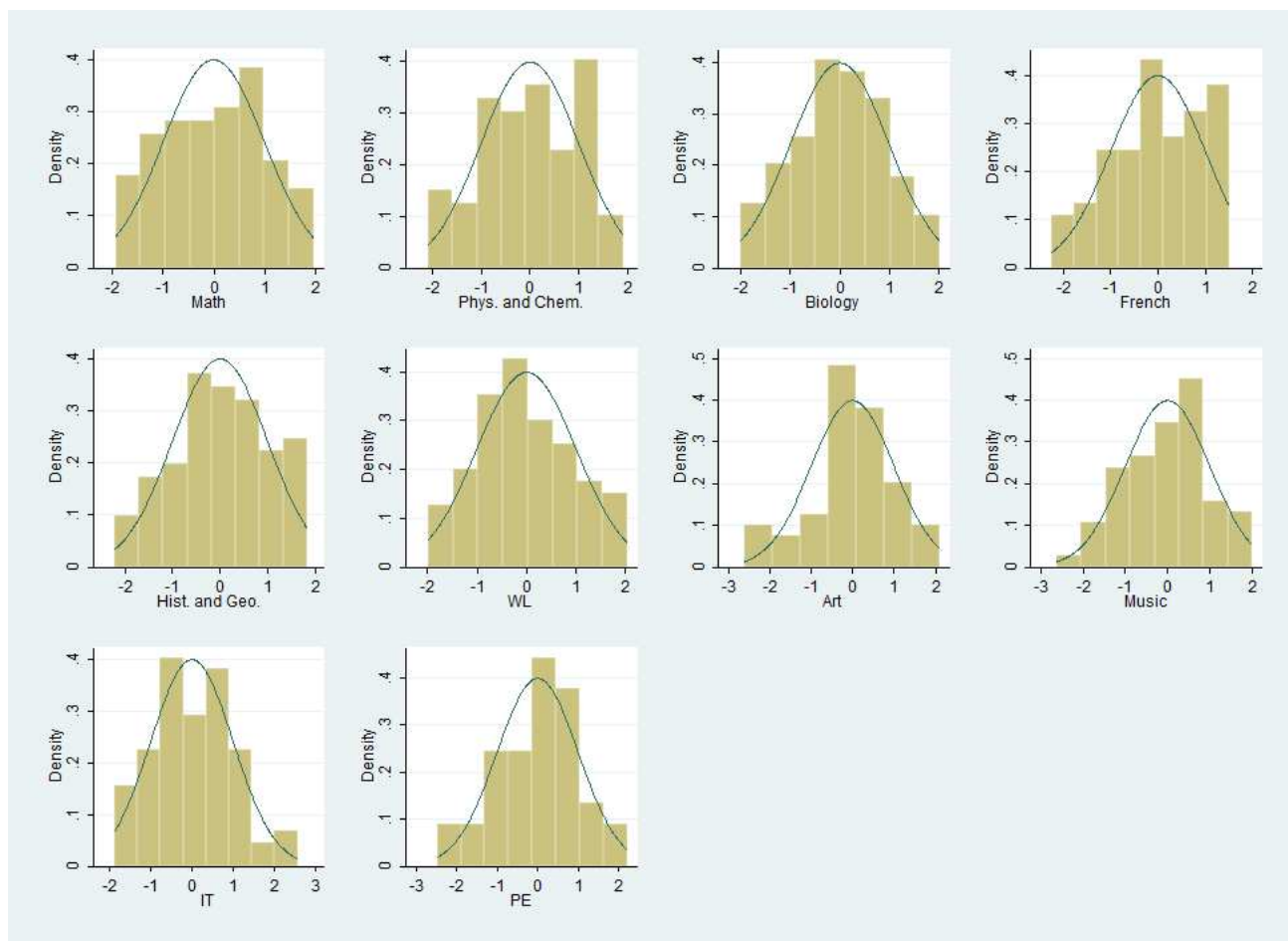


Figure 3.3: Distribution of the standardized grades in each subject

At the end of their 9th grade, the pupils take the middle school final exam ("le Brevet des Collèges") that delivers a diploma. It contains three tests: one in math, one in French and one in history and geography. The success at this exam is partly based on the grades obtained to these written exams and partly on the continuous assessment of 9th grade. Each of these evaluations weight 50% in the final grade that allows getting the diploma.

Nowadays, this exam has become informal and it is possible to pursue to high school

even though one has not passed it. Nevertheless, most of the pupils still take it. We obtained the information if they passed it or not for 125 pupils¹¹, and whether they got a distinction, but we did not get their exam grades. This exam is anonymous and pupils are graded by external teachers who do not know about the name or gender of the pupil.

Among pupils for which we will estimate the grades in the following section, their success rate is equal to 71.6%. It is a fairly low success rate as the national success rate for this diploma has been between 80% and 85%¹² these last ten year¹³. 64.6% of them got a distinction. There are three distinctions for the French diploma: the lowest distinction (cum laude) is awarded if one reaches an average grade between 12/20 and 14/20 (magna cum laude), the second one for an average grade between 14/20 and 16/20 (summa cum laude), and the highest one for an average grade higher than 16/20.

3. Results

This paper aims at explaining schooling achievement with creativity scores and personality traits. First, we will shed lights on who might be the creative pupils. Then, we will estimate the standardized grades in each subject with OLS regression. As we went three times to the middle school, we face a problem of selection bias. Indeed, pupils who attended the three sessions are pupils who do not skip classes. They can be considered as more conscientious and regular¹⁴. Over the 99 pupils who attended

11. Even though we did not get some pupils' grades' report, they still took the exam. The mean average one obtains is on a 40 points basis but that can be easily scaled to a 20 points scale. To pass this exam, one should get an average grade of 20/40, based on the continuous assessment and exam grades. Getting a 0/40 in one of the subject exam test yields immediate failing.

12. This information is available on <http://www.education.gouv.fr/cid59753/diplome-national-du-brevet.html>

13. This is an indication of the school level being rather low.

14. However a two-tailed t-test on the score of conscientiousness between those who attended the creativity sessions and those who did not yields a p-value=0.39.

the BB5 session, 81 attended the creativity sessions. Overall, the pupils who attended the three sessions have significantly higher grades compared to those who did not¹⁵.

3.1. Who are the creative pupils?

Before going into greater estimations on the relationship between creativity and schooling achievement, we try in this subsection to have a broader idea of who the creative pupils are.

Some personality traits are usually associated with creative potential such as openness or individualism. Table 3.3 reports the correlations coefficients between the creativity variables and the personality traits.

	GIT	VIT	GDT	VDT
Agreeableness	0.2910 (0.0113)	0.2540** (0.028)	0.1023 (0.383)	0.2003* (0.085)
Conscientiousness	0.0129 (0.912)	0.1431 (0.221)	-0.0436 (0.710)	-0.0118 (0.920)
Extraversion	-0.0084 (0.943)	0.1821 (0.118)	0.0991 (0.397)	0.2532*** (0.028)
Emotional Stability	0.1520 (0.193)	0.2052* (0.077)	0.1328 (0.256)	0.1506 (0.197)
Openness	0.2050* (0.078)	0.3338*** (0.003)	0.1161 (0.321)	0.3225*** (0.005)

Table 3.3: Correlation between the BB5 personality traits and the creativity scores (p-values in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

The correlations coefficients are globally in line with the results of Barbot (2012). Indeed, he mainly finds a high correlation (significant at a 5% level) between openness

15. Two tailed t-tests on the standardized grades for each subject yield a p -value < 0.1 but for biology and sports.

and GDT, VDT and VIT, and also a correlation between extraversion and GDT¹⁶. In our case, openness is highly correlated with both of the verbal tasks and extraversion is correlated with VDT. We also find a weak correlation between agreeableness and verbal tasks, but not as significant as the other correlation coefficients.

Therefore, openness seems to be more correlated to all subtypes of creativity compared to the other personality traits. This is consistent with the result that openness facilitates the use of divergent thinking for more creative production (McCrae and Costa, 1987). We can see that it is also the case for integrative thinking. Creativity can thus be considered as a cognitive ability that is rather poorly correlated to personality traits. They are independent variables suspected here to play a role in schooling success.

As we saw before, higher scores are found for integrative thinking (both graphic and verbal). We look at gender in order to see if girls and boys exhibit the same degree of creativity. Figure 3.4 displays the distribution of each creativity scores by gender.

Overall, creativity scores' distributions are more shifted on the right for girls which would imply higher creativity scores than for boys. This is less clear for the GIT dimension where both distributions look alike (except the peak for girls where 30% of the sample has a 0 standardized score of GIT). In fact, two-tailed t-tests inform that scores in GIT are not significantly different between boys and girls ($p=0.83$), which is also the case for the GDT scores ($p=0.14$). However, for both verbal tasks, girls have a higher score than boys (for the VIT score the p -value= 0.001 and for the VDT score the p -value= 0.06).

Baer and Kaufman (2008) review the topic of gender differences in terms of creativity. Among the different studies on that topic, results are not always consistent mainly because of the multiple creativity measures used. Still, a general trend seems

16. His study is based on 607 teenagers.

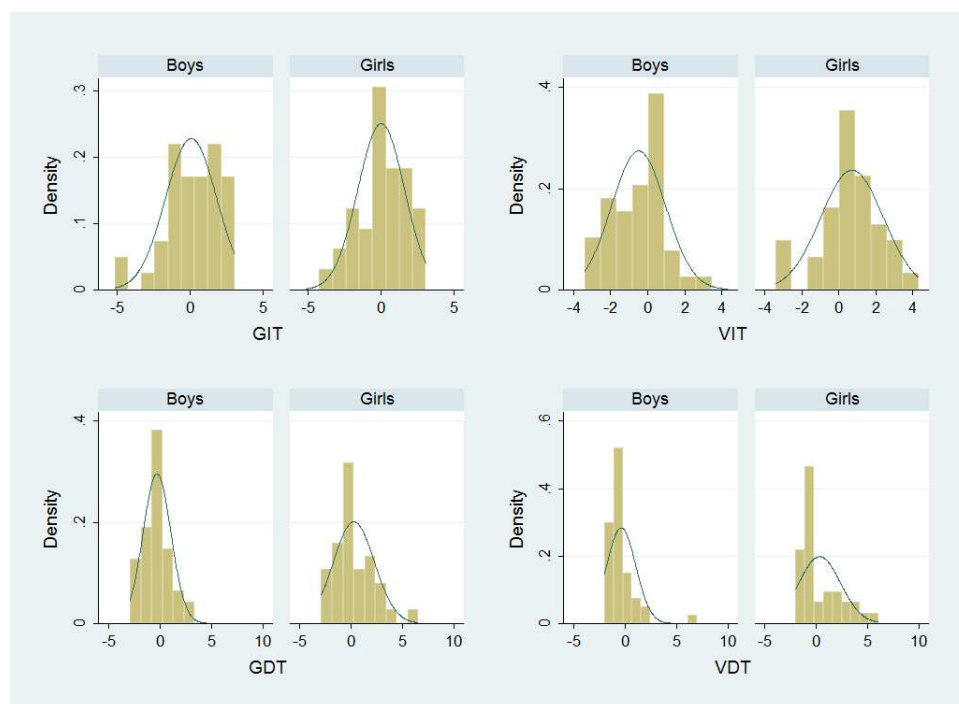


Figure 3.4: Creativity scores by gender

to be that women and girls have higher creativity scores than men and boys.

3.2. Estimation of schooling achievement

In this section, we estimate the standardized average grade in each subject on the whole 9th grade year¹⁷ (three trimesters). We do so, first without class fixed effects, and second by integrating them in the estimation. The underlying research question is whether creativity scores and other individual characteristics have an impact on grades.

As for the independent variables, we use the five personality traits scores, the creativity scores and demographics such as gender and age as well as a dichotomous variable indicating inconsistency during the lottery task.

17. There are no significant trend on the year meaning that pupils globally exhibit constant grades during the school year. That is why we can compute the year average grade.

The whole equation we estimate is the following:

$$Grade_{ijc} = \beta C_i + \gamma P_i + \lambda X_i + \alpha_c + \epsilon_{ij}$$

for a pupil i and $j = \{math, French, history and geography etc...\}$ corresponds to the different subjects, c the class indicator. C_i is the vector of the four individual creativity scores, P_i the vector of the five personality traits, X_i the other observed covariates (gender, age, inconsistency) and α_c the class fixed effect (in order to control for any unobserved heterogeneity coming from the class). These estimations are reported in tables 3.4, 3.5, 3.6.

Globally, the verbal divergent thinking estimates are negative for all subjects and are significantly associated to math, biology and French, and close to significance for physics. Divergent thinking was the first dimension studied by psychologist to evaluate creativity potential. It is considered as essential for creative processes. It can be seen as the basic ability that allows to create alternative solutions to a given problem. Estimations inform on the fact that, in this school, the higher the VDT score, the worse the pupil's grade. This would suggest that the grading system does not encourage the use of divergent thinking abilities that are fluency, flexibility and originality. It might mean that the school's teachers evaluate pupils in a quite conservative way, which supports the criticism that can sometimes be made to the French educational system.

Surprisingly, the other dimensions of creativity have no effect on our variable of interest except graphical integrative thinking which is positively and significantly associated with music's and IT's grade. The non significant effect of the graphic dimensions may be due to the fact that at this age, pupils have to invest more in verbal domains rather than in graphic domain. Even when a pupils has a high score in graphic thinkings, it might just not be needed to perform well at school.

In terms of personality, openness is the main trait that is positively associated with almost all the grades. This is consistent with the literature on this trait. Previous research has found that openness which reflects curiosity, imagination, unconventionality

VARIABLES	Math (1)	Math (2)	Phy. & Che. (3)	Phy. & Che. (4)	Biology (5)	Biology (6)
Girl	1.071*** (0.218)	0.725*** (0.252)	0.916*** (0.215)	0.696*** (0.223)	1.141*** (0.209)	0.682*** (0.221)
Age	-0.005 (0.162)	-0.041 (0.185)	-0.110 (0.199)	-0.127 (0.204)	0.052 (0.215)	-0.028 (0.240)
GIT	0.041 (0.077)	0.055 (0.077)	0.012 (0.087)	0.036 (0.095)	0.073 (0.091)	0.077 (0.094)
VIT	-0.072 (0.070)	-0.001 (0.075)	-0.007 (0.071)	0.026 (0.080)	-0.085 (0.074)	0.028 (0.076)
GDT	0.005 (0.066)	-0.055 (0.062)	-0.036 (0.069)	-0.071 (0.072)	0.065 (0.067)	-0.033 (0.061)
VDT	-0.130** (0.063)	-0.104* (0.064)	-0.108* (0.063)	-0.087 (0.069)	-0.163*** (0.061)	-0.164*** (0.053)
Agreeableness	0.001 (0.014)	0.012 (0.014)	0.001 (0.015)	0.009 (0.016)	0.016 (0.016)	0.027 (0.017)
Conscientiousness	0.009 (0.013)	0.019* (0.011)	0.005 (0.014)	0.012 (0.013)	-0.002 (0.014)	0.005 (0.012)
Extroversion	-0.020* (0.011)	-0.012 (0.011)	-0.014 (0.012)	-0.011 (0.014)	-0.010 (0.010)	0.003 (0.010)
Emotional Stability	0.004 (0.011)	-0.014 (0.012)	-0.005 (0.013)	-0.014 (0.015)	0.004 (0.013)	-0.018 (0.013)
Openness	0.015 (0.015)	0.032** (0.015)	0.033** (0.016)	0.038** (0.018)	0.028* (0.014)	0.046*** (0.016)
Incons.	-0.526** (0.236)	-0.515** (0.220)	-0.528** (0.213)	-0.503** (0.208)	-0.335 (0.236)	-0.350* (0.201)
Constant	-0.844 (2.727)	-1.624 (2.985)	0.087 (3.471)	-0.468 (3.557)	-3.443 (4.078)	-3.120 (4.393)
Class FE	NO	YES	NO	YES	NO	YES
Observations	74	74	73	73	72	72
R-squared	0.312	0.479	0.319	0.414	0.406	0.580

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All regressions are estimated using OLS, with robust standard errors. Girl is a dichotomous. Age is a continuous variable. Incons. is a dichotomous variable equal to 1 if the pupil exhibits inconsistency regarding the risk aversion measure, 0 otherwise.

Table 3.4: OLS estimation on scientific subjects' grades (standardized)

(also sometimes called intellect or openness to new experience) is correlated with intelligence scores. A higher score of conscientiousness also increases the grades in physics (when no class fixed effects are included), music and IT. We would have expected a higher predictive power of this trait on other grades (estimates are positive but insignif-

VARIABLES	French (1)	French (2)	Hist. & Geo. (3)	Hist. & Geo. (4)	WL (5)	WL (6)
Girl	1.217*** (0.216)	0.906*** (0.226)	0.961*** (0.237)	0.726** (0.314)	0.634** (0.246)	0.295 (0.293)
Age	0.152 (0.184)	0.091 (0.180)	-0.023 (0.175)	-0.109 (0.215)	-0.010 (0.211)	-0.025 (0.222)
GIT	0.046 (0.086)	0.082 (0.086)	0.043 (0.083)	0.054 (0.090)	0.006 (0.098)	0.013 (0.101)
VIT	-0.036 (0.073)	0.057 (0.071)	-0.060 (0.074)	0.003 (0.085)	0.045 (0.085)	0.114 (0.078)
GDT	0.098 (0.060)	0.048 (0.063)	0.009 (0.068)	-0.030 (0.076)	0.019 (0.078)	-0.041 (0.087)
VDT	-0.142*** (0.053)	-0.117** (0.047)	-0.073 (0.067)	-0.049 (0.067)	-0.025 (0.074)	-0.008 (0.065)
Agreeableness	0.014 (0.015)	0.025* (0.013)	-0.012 (0.014)	-0.006 (0.016)	0.009 (0.018)	0.019 (0.019)
Conscientiousness	0.014 (0.013)	0.018 (0.011)	0.018 (0.013)	0.021 (0.013)	0.013 (0.014)	0.023* (0.013)
Extroversion	-0.006 (0.012)	0.009 (0.011)	-0.009 (0.014)	-0.000 (0.015)	-0.007 (0.015)	0.001 (0.016)
Emotional Stability	-0.012 (0.012)	-0.032** (0.013)	-0.000 (0.014)	-0.013 (0.016)	-0.013 (0.014)	-0.030* (0.015)
Openness	0.013 (0.014)	0.033** (0.015)	0.019 (0.020)	0.032 (0.021)	0.015 (0.020)	0.033 (0.021)
Incons.	-0.052 (0.251)	-0.002 (0.220)	-0.159 (0.254)	-0.133 (0.259)	-0.221 (0.278)	-0.240 (0.258)
Constant	-4.198 (3.378)	-4.838 (2.964)	-0.994 (2.965)	-0.676 (3.594)	-1.161 (3.809)	-2.184 (3.986)
Class FE	NO	YES	NO	YES	NO	YES
Observations	73	73	74	74	73	73
R-squared	0.434	0.599	0.260	0.357	0.210	0.334

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: OLS estimation on humanity subjects' grades (standardized)

icant) as this trait includes control, organization or self-efficacy. [Blickle \(1996\)](#) shows that conscientiousness and openness have an effect on learning strategies¹⁸. [Goff and Ackerman \(1992\)](#) and [Ashton et al. \(2000\)](#) also show that openness and intellectual ability are positively correlated.

18. They use scales to evaluate various strategies such as cognitive learning strategies, or resource-related leaning strategies.

VARIABLES	Art (1)	Art (2)	Music (3)	Music (4)	IT (5)	IT (6)	PE (7)	PE (8)
Girl	0.568** (0.234)	0.441** (0.203)	0.884*** (0.263)	0.943*** (0.270)	0.562** (0.278)	0.169 (0.278)	-0.381 (0.261)	-0.551** (0.253)
Age	-0.419** (0.190)	-0.262 (0.156)	-0.164 (0.238)	-0.078 (0.208)	0.087 (0.148)	0.082 (0.159)	0.188 (0.212)	0.172 (0.201)
GIT	0.019 (0.062)	0.022 (0.051)	0.143** (0.069)	0.121* (0.069)	0.121* (0.066)	0.120* (0.068)	0.029 (0.075)	0.023 (0.077)
VIT	0.087 (0.103)	0.037 (0.085)	0.110 (0.076)	0.110 (0.083)	-0.003 (0.080)	0.069 (0.068)	0.104 (0.075)	0.144** (0.062)
GDT	0.046 (0.066)	-0.020 (0.061)	-0.028 (0.065)	-0.018 (0.078)	-0.009 (0.072)	-0.091 (0.074)	-0.119 (0.085)	-0.153** (0.076)
VDT	-0.110 (0.070)	-0.086 (0.091)	-0.080 (0.052)	-0.042 (0.049)	-0.109 (0.067)	-0.105 (0.073)	-0.072 (0.082)	-0.030 (0.069)
Agreeableness	-0.015 (0.016)	-0.003 (0.012)	0.011 (0.017)	0.017 (0.017)	-0.006 (0.014)	0.004 (0.014)	-0.014 (0.013)	-0.006 (0.012)
Conscientiousness	-0.022 (0.018)	0.007 (0.014)	0.028** (0.011)	0.035*** (0.012)	0.012 (0.012)	0.022** (0.010)	0.013 (0.013)	0.019* (0.010)
Extroversion	-0.024 (0.017)	-0.027 (0.017)	0.006 (0.011)	0.001 (0.011)	-0.038*** (0.013)	-0.025* (0.013)	0.008 (0.014)	0.012 (0.014)
Emotional Stability	0.015 (0.020)	-0.007 (0.018)	-0.021* (0.013)	-0.032** (0.014)	0.012 (0.012)	-0.012 (0.013)	0.022 (0.013)	0.002 (0.014)
Openness	0.035** (0.016)	0.042** (0.018)	0.040** (0.015)	0.062*** (0.016)	0.023 (0.018)	0.043** (0.020)	-0.001 (0.016)	0.015 (0.015)
Incons.	-0.223 (0.328)	-0.193 (0.240)	0.015 (0.213)	0.140 (0.201)	-0.316 (0.251)	-0.317 (0.250)	-0.125 (0.266)	-0.015 (0.224)
Constant	6.466** (2.971)	3.307 (2.593)	-2.172 (3.830)	-4.694 (3.095)	-1.591 (2.766)	-2.498 (2.904)	-4.267 (2.997)	-4.820* (2.663)
Class FE	NO	YES	NO	YES	NO	YES		
Observations	53	53	60	60	74	74	72	72
R-squared	0.396	0.641	0.604	0.698	0.288	0.467	0.218	0.405

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.6: OLS estimation on art, music and IT and sport grades (standardized)

Having a higher score of openness is hence rewarded by better grades in our sample. But, having a high score of VDT is not, even though openness is positively correlated with this dimension. This might be explained by the fact that schools ask pupils to manifest curiosity, imagination and a taste for novelty but to stay in compliance with specific learning methods. This is incompatible with divergent thinking as it relies on the ability to manage to consider a problem from different and new perspectives.

Table 3.8 in the appendix presents the distribution of grades for each subject between girls and boys. We can see that the distribution of girls' grades are more shifted to the right than the boys' grades. Indeed, the estimates of the girl dummy variable in tables 3.4, 3.5 and 3.6 the are positive and highly significant for every subject but PE. This may be not surprising as girls are considered as more attentive, focused and self-disciplined (see Costa Jr, Terracciano and McCrae (2001), Rubinstein (2005)). In our sample, girls are significantly more agreeable, more open and slightly more conscientious (see table 4.1 in the annex for details). It is usually asserted that girls have better grades because they are more conscientious. Even though we control for these traits, we still find a great significant differences between boys' and girls' grades. Hence, the gender itself hints a mechanism not captured by all the variables we control for. Research paper investigate the question of stereotype threat in school. By comparing non-anonymous vs. anonymous exam, they test the assumption that girls are either positively or negatively discriminated by the teacher. Both Lavy (2008) and Falch and Naper (2013) finds as well that girls outperform boys in almost all the subjects and suggests that the bias they found against male students is the results of teachers' behavior¹⁹. One reason for this result could be that girls and boys invest differently in the subject according to the teacher's gender. Indeed, the way students perceive their teacher seem to have an impact on their motivation and performance (Maehr and Midgley, 1991, Meece, Glienke and Burg, 2006). A paper by Dee (2007) shows that teachers exhibit better perception when students are from their own gender.

Ouazad and Page (2013) show that in school, boys invest less when they are graded by a female teacher, and girls invest more when they are graded by a male teacher. Our sample size is rather small and comes from only one school but we have information about the gender of the teachers. Even though there are for instance three different math teachers, they are all women. In physics there are two different teachers, but

19. The first study is run on Israeli's data, while the latter was done in Norway.

one of them is a man and teaches to four classes (over the six). In any cases, girls have higher grades. At least on our small sample, boys and girls do not seem to be differently graded according to the gender of the teacher.

The assumption of favoring girls can be tested by looking at the success probability of the national and anonymous exam, estimated in the following section.

3.3. Estimation of the probability of passing the middle school final exam

We are interested in explaining the probability of succeeding this exam. We use the same independent variables as in the previous section.

Table 3.7 reports a probit on the dummy variable of passing (=1) or not (=0) the exam. As age seems to matter here (which is not the case in our previous estimations on grades) we create a dummy variable equal to 1 if the pupil is older than 14. The regular age in 9th grade is 14 or 15 years old. Being older can either mean being born at the very beginning of the year or it can also mean having repeated the grade, and being younger skipping one grade²⁰.

Passing this exam also depends on a continuous assessment based on 9th grade's scores. This data was not available. In order to take this into account, we assume that having good grades during the year would increase the probability of passing the exam. Hence, a rather good control for this is to add the average grades during the year. We choose to add those in scientific and humanities as a proxy for the continuous achievement. Pupils with a higher average grade in sciences and humanities will have a higher grade on their continuous assessment.

The interesting result given by this table is that variables associated with the probability of passing the exam are not the same as for getting better grades during the school year. Column (1) reports the estimates of the probit when we do not take into

20. We did not have access to their exact date of birth but the sessions during which we asked their age were run in January 24th and 25th. The distribution of age is the following: 2.7% of our sample is 13, 64.0% is 14, 28.0% is 15 and 5.3% is 16. The 16 year old and certainly most of the 15 year old pupils are likely to be repeaters.

VARIABLES	Passed (1)	Passed (2)
Girl	-0.679 (0.501)	-1.209* (0.641)
Age>14	-1.715*** (0.360)	-1.465*** (0.518)
Sc. grades	1.784*** (0.467)	2.343*** (0.692)
Human. grades	-0.358 (0.509)	-0.559 (0.712)
GIT		0.366 (0.288)
VIT		0.809*** (0.238)
GDT		-0.159 (0.170)
VDT		0.188 (0.196)
Observations	85	75

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.7: Probit on the passing the final exam

account creativity nor personality. Higher grades in science predicts positively and significantly the probability of passing the exam. However humanity grades have no impact even though there is one test in French and one in history & geography.

Columns (2) introduces creativity scores²¹. Surprisingly, the creativity dimension that has a positive impact here is the verbal integrative thinking. As mentioned before, we expected that this creativity dimension to be positively associated with schooling grades. However, we found previously that VIT had no significant effect on grades. Moreover, the VDT estimate is here not significant but was negative and significant in the previous analysis on schooling grades. This raises a question on the abilities

21. No effect of personality was found and did not change qualitatively and quantitatively the estimates of the other covariates. We therefore did not report the estimates.

needed for achievement during the year and during an anonymous national exam. Does schooling limit the use of VIT during the year even though it is necessary in order to succeed an exam (where pupils must have a global knowledge based on many years of studies)?

Concerning girls, another interesting results emerges. The estimate of the girl's dummy variable is negative and significant (at a 10% level) meaning that, all things being equal, girls have a lower probability to pass the exam than boys. Nevertheless, we showed in the previous subsections that they had a higher score of VIT and had better grades. Indeed, 63% of boys passed the exam whereas 71% of girls did. The negative estimates indicates that if we compared a boy and a girl having the same level of ability, creativity and personality, boys might succeed better. This will be later discussed.

Finally, being older than 14 decreases the probability of success suggesting that repeating a grade might not allow passing this exam.

4. Discussion

In France, orientation choices, which occur at the end of middle or high schools, are made before taking the national exams (Brevet des Collèges and Baccalauréat). They are based on the schooling grades and not on the exams' scores²². Succeeding the national exams is only needed in order to be allowed to pass to the next educational level²³. Hence, explaining schooling grades remains a first issue to consider, followed by the exam success.

Openness and conscientiousness seem to be the main predictor of better schooling grades. As openness contributes to creativity, we could think that originality is re-

22. These decisions might be however influenced by the exam's results ex post, but this concerns a marginal part of the pupils.

23. In fact, it was mandatory to pass the middle school exam, and it is still mandatory to pass the high school exam in order to be able to reach to higher educational levels.

warded by better grades. Indeed, teenagers with a high score of openness are curious and new experience seekers. They can be considered as eager to learn, artistic and imaginative. Nevertheless, we find that a higher score of verbal divergent thinking decreases grades. This raises the question of which type of creativity is beneficial to the pupils in school. From our study, it seems that open pupils have better grades if they manage to be in line with the standard learning strategies. Would the French educational system be less considered as conventional if new learning tools would insist on the development of creativity? This remains an open question and further studies are needed, especially if creativity represents really an important skill when entering the labor market.

Concerning the gender effect, we observe that girls have better grades during the school year, but do not necessarily succeed better at the exam when compared to a boy having the same level of abilities. We draw different explanations that we are not able to confirm with the lack of available data but that can be further explored: girls can choke under the pressure of this national exam more than boys and thus succeed less. Another possible explanation is that girls are maybe more hard-working during the whole year and more conscientious leading to better results during the year. But, before an exam, boys may better perform when studying for the exam and would then outperform girls. Lastly, our result could also be explained by the idea developed by [Lavy \(2008\)](#) and [Falch and Naper \(2013\)](#) i.e. that teachers give higher grades to pupils when they know they are girls, implying a negative bias towards boys. In anonymous exam, this positive discrimination cannot be possible.

5. Conclusion

This paper establishes a link between academic achievement, creativity and personality traits. Considering the impact of creativity on schooling achievement has been rarely studied in the growing economic literature of the analysis of non-cognitive abil-

ities and educational outcomes. We went on the field, in a middle school, where we were able to measure creativity potential of 9th graders and obtained four scores of creativity subtypes: verbal divergent thinking, graphical divergent thinking, verbal integrative thinking and graphical divergent thinking. One has to note that our sample comes from low socio-economics status²⁴, that is why one must be careful when extending the our results.

Academic achievement was assessed with the report of grades of their 9th grade for the mandatory courses. We could have expected that creativity might play a role in having better grades, at least in the creative subjects such as arts and music. We find that the different subtypes of creativity do not affect grades in the same way. Verbal divergent thinking predicts significantly and negatively grades' subjects. This result suggests that this is not a skill required to succeed at school. However, it is required in many innovative domains as in R&D. We are also able to estimate the probability of passing the "Brevet". In contrast, a higher score of integrative thinking increases the probability, but being a girl does not.

We also confirm previous work on the relationship between personality traits and schooling achievement: conscientiousness and openness are the main traits that positively influences grades.

From these results, we are able to present some recommendations as for the future studies on schooling achievement. One recommendation relates to taking into account the specificities of the different learning domains. Ability and success cannot always be considered globally. An IQ test, or other global math or word tasks test are not able to highlight this specificity. As we saw previously, different individual variables impact schooling grades in different ways.

It might be of interest to get a deeper understanding of the role of creativity at

24. We asked pupils the socio-professional category of their parents as well as their exact job. We were then able to establish the proportion of high qualified mothers and fathers in our sample which is respectively 35.4% and 29.1%.

school, at different levels of education and its implication on the labor market. Second, gender is a key variable that might imply other underlying behaviors or mechanisms.

Further development could require replicating the same study in different schools, pedagogies, with different types of pupils (different social background, neighborhood), and even abroad where different educational systems are assumed to emphasise more on creativity. We would then be more able to describe profiles of creative pupils: who are the most creative pupils? Is there a relationship between being creative and drop outs? Are the creative pupils have different grades profiles than the other pupils? These are the remaining questions we would like to answer.

C. Appendix

C.1. Figures



Figure 3.5: Two examples of GIT task.

La goutte d'eau.

Kaylinn regardait tranquillement une émission à la télé. Tout d'un coup, son frère Kevin lui lance de l'eau sur le visage. alors, Kaylinn se lève, puis lui court après dans toute la maison, Kevin se cache, va dans la cuisine et prend un oeuf. Bien caché, il attend d'avoir Kaylinn dans sa ligne de tir. Quand il la voit enfin, il lui jette l'oeuf dessus. S'en est trop, "C'est la goutte d'eau qui fait déborder le vase," dit elle. Elle attend alors que Kevin sorte de sa cachette, mais en vain. Le lendemain matin, la jeune fille voit Kevin dans son lit. Elle va chercher de l'eau, puis à son tour lui jette de l'eau. Kevin sursaute, et tous les deux se mettent à rire.

La goutte d'eau.

C'était un jour de pluie, et je voulais aller dehors pour m'amuser dans boue, mais maman ne voulait pas. Mais moi je suis têtue, alors je me précipite dehors pour m'amuser avec mes copains. Plus le temps ^{passé} plus je m'amuse, mais la pluie ne cessait de tomber. Au bout d'un moment la pluie était tellement forte qu'elle martelait le ~~de~~ ^{dessus} de la terre. Lorsque tous mes copains étaient rentrés je me retrouvais seul dans ~~la~~ ^{ma} chambre sous une pluie battante. Quand tout à coup j'entendis un énorme bruit, je lève la tête lorsque j'aperçois une énorme goutte, plus grosse que moi qui arrivait à toute vitesse pour me percuter... et là, je me réveille en sursaut dans mon lit trempé de sueur. Ouff! ce n'était qu'un rêve! À l'avenir j'écouterais maman!

Fin

Figure 3.6: Two examples of VIT task: writing a story based on the title "Drop of Water"

C.2. Tables

- Claude vit un éléphant marcher et écraser tout ce qu'il y avait sur son passage, d'un seul coup, il arrêta de jouer et rentra chez lui.
- Claude vit un immeuble s'effondrer
- Claude vit que ce bruit n'était que celui d'une jeune fille ne sachant pas chanter, puis il se moqua d'elle.
- Claude vit un tracteur rouler près du parc.
- Il tanna sa tête, et vit le sol fissuré
- Claude n'eut pas le temps de tourner la tête, qu'il était poursuivi par un chameau
- Claude rentra chez lui
- Mais il continua à jouer sans s'en occuper
- Il vit qu'un enfant avait cassé une des installations de l'air de jeux.
- Il se retourna, et vit, à terre, un chat qui lui semblait bien mal en point, il avait une voix bizarre.

Un après Claude jouait au Parc lorsque il entend un bruit..

- ① Il se précipita aperçut derrière un buisson, il n'osait même ^{pas} ouvrir les yeux. Claude prit son courage à deux mains puis sorti de sa cachette lorsque il aperçut que ce n'était petit chaton, il se senti ridicule.
- ② Claude très curieux s'empressa d'aller voir ce qu'il se passait, mais le bruit se déplaçait de plus en plus vite. Claude le suivait toujours sans savoir vraiment où il allait. Au bout d'un moment il arriva dans les bois et commença à avoir peur, quand tout à coup 3 jeunes ~~un~~ garçon sortis des buisson. Mais ce n'était rien, juste ses amis qui voulait lui faire une peur.
- ③

Figure 3.7: Two examples of VDT task: writing as many ends of story as possible

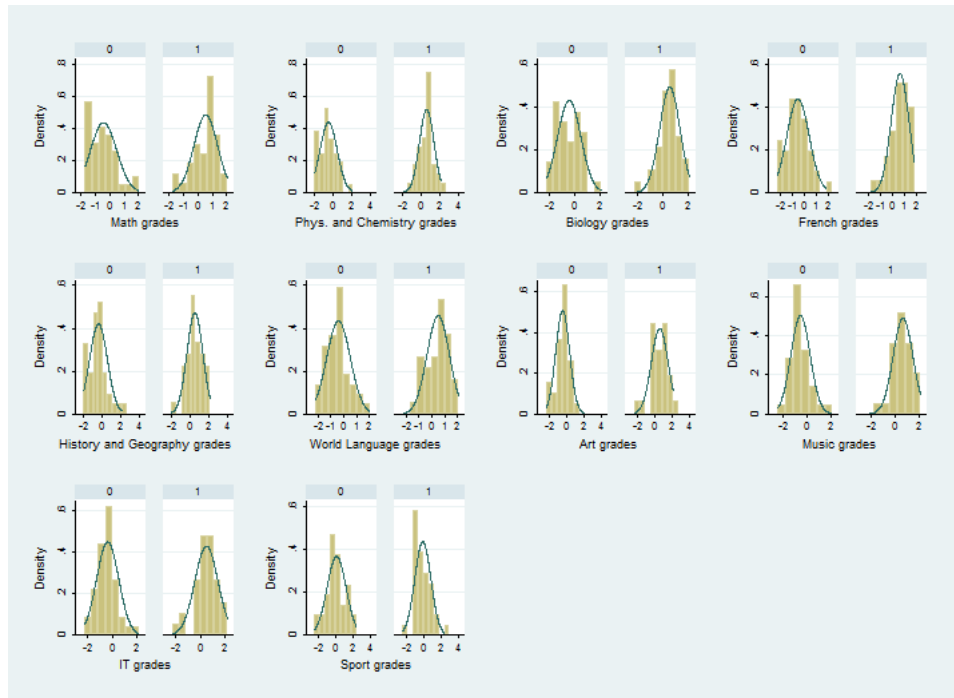


Figure 3.8: Distribution of standardized grades per subject, by gender (0 for boys, 1 for girls)

	Girls	Boys	T-test p-value
Agreeableness	67.7	62.4	0.00
Conscientiousness	64.4	60.9	0.09
Extraversion	64.3	62.5	0.34
Emotional Stability	64.5	64.4	0.99
Openness	62.7	59.4	0.04
	n=42	n=49	

Table 3.8: BB5 scores by gender

Variable	Obs	Mean	Std. Dev.	Min	Max	Variable	Obs	Mean	Std. Dev.	Min	Max
AvGr_MathA	16	11.7	4.4	4.2	16.5	AVaGr_WLA	16	11.9	4.1	4.8	19.0
AvGr_MathB	18	12.0	3.5	4.3	18.3	AvGr_WLB	18	11.9	4.2	3.5	17.5
AvGr_MathC	16	11.7	2.9	6.5	16.8	AvGr_WLC	16	11.0	2.8	6.7	15.8
AvGr_MathD	10	13.4	4.4	5.5	17.8	AvGr_WLD	10	11.8	4.1	4.3	16.8
AvGr_MathE	20	9.3	4.0	4.3	17.7	AvGr_WLE	20	9.9	3.5	3.7	18.5
AvGr_MathF	17	10.3	4.0	4.2	19.0	AvGr_WLF	16	10.3	4.5	3.8	18.5
AvGr_Phys.A	16	11.2	4.3	2.7	17.2	AvGr_ArtsA	1	15.5	,	15.5	15.5
AvGr_Phys.B	17	13.0	3.5	7.2	17.2	AvGr_ArtsB	17	15.9	1.9	13.5	19.5
AvGr_Phys.C	15	12.8	3.0	4.8	17.5	AvGr_ArtsC	15	13.0	2.1	6.5	15.7
AvGr_Phys.D	10	12.0	3.5	6.2	15.7	AvGr_ArtsD	1	13.2	,	13.2	13.2
AvGr_Phys.E	20	10.4	4.4	3.8	18.7	AvGr_ArtsE	20	11.5	3.1	5.8	16.7
AvGr_Phys.F	17	10.8	4.0	4.5	19.0	AvGr_ArtsF	16	12.6	2.3	6.5	17.7
AvGr_BioA	15	12.4	2.8	5.8	17.3	AvGr_MusicA	14	13.7	2.8	9.5	18.5
AvGr_BioB	18	10.4	4.3	1.2	17.0	AvGr_MusicB	15	15.0	2.5	9.7	19.7
AvGr_BioC	16	10.8	2.5	5.5	15.3	AvGr_MusicC	15	14.3	2.7	6.7	18.0
AvGr_BioD	10	12.3	1.9	9.3	14.8	AvGr_MusicD	1	15.0	,	15.0	15.0
AvGr_BioE	20	8.6	3.5	4.2	15.0	AvGr_MusicE	20	12.9	3.0	7.8	19.3
AvGr_BioF	15	9.3	3.3	4.7	16.8	AvGr_MusicF	15	13.6	3.0	8.5	19.5
AvGr_FrenchA	16	10.8	3.4	4.8	15.3	AvGr_ITA	16	13.4	1.9	9.2	15.5
AvGr_FrenchB	16	10.7	3.3	3.8	15.7	AvGr_ITB	17	12.7	1.3	10.5	15.0
AvGr_FrenchC	16	11.5	2.6	4.7	15.3	AvGr_ITC	16	13.0	1.7	10.2	17.2
AvGr_FrenchD	10	13.8	1.6	11.2	15.3	AvGr_ITD	10	13.0	2.5	9.7	16.7
AvGr_FrenchE	20	8.9	2.7	4.7	14.7	AvGr_ITE	20	11.7	1.5	9.5	14.8
AvGr_FrenchF	17	9.4	3.8	1.8	15.5	AvGr_ITF	17	12.3	1.6	9.7	15.8
AvGr_H&GA	16	11.5	2.9	7.2	15.7	AvGr_PEA	16	13.2	1.7	10.7	15.8
AvGr_H&GB	18	10.9	3.3	1.2	14.7	AvGr_PEB	16	13.5	1.7	9.3	16.2
AvGr_H&GC	16	12.6	3.2	6.0	17.2	AvGr_PEC	16	14.0	2.7	7.3	18.0
AvGr_H&GD	10	13.2	2.6	9.7	17.0	AvGr_PED	8	13.8	2.6	11.2	18.7
AvGr_H&GE	20	9.8	3.2	4.3	16.5	AvGr_PEE	18	11.5	2.4	8.0	15.3
AvGr_H&GF	17	10.0	4.2	1.8	16.7	AvGr_PEF	17	14.1	2.0	10.3	17.3

Table 3.9: Descriptive statistics on grades in each subject and class (A to F), on a 20 points basis

CHAPTER 4

A HOLT & LAURY MEASURE OF RISK
AVERSION ON TEENAGERS

1. Introduction

The way young economic agents make their decisions may have a crucial impact on performance of economic policies. Teenagers differ from adults in the way they behave and make their decisions. The reason for that is partly biological because the brain continues to mature and to develop at this age. For example, the area of the brain that controls reasoning and help people think before they act (the frontal cortex) develops at a later age. Hence, because of this stage of brain development, adolescents are for example more likely to act on impulse, get involved in fights or engage in dangerous risky behaviors. Apart from a biological explanation, their higher propensity to take risks can also be explained by their vulnerability to social pressures (Steinberg, 2004), because they are more sensitive to emotional stimuli and lack of cognitive control. Peers presence significantly increases risky behaviors (Gardner and Steinberg, 2005). Berndt (1979) and Brown, Clasen and Eicher (1986) identified a linear relationship between age and peer influence. Moreover, risk taking represents a way for the teenager to develop his social and psychological abilities (independence, autonomy and self-regulation) in an optimal and efficient manner (Silbereisen and Noack, 1988). This does not mean they are not able to make good or thoughtful decisions, but it suggests that if one better understands how these decisions are made, policies aimed at children and teenagers may become more efficient and decisions more optimal. In fact, psychological studies show that 15-year-olds are no worse than adults in understanding risk (Reyna and Farley, 2006) and increasing the salience of the risks associated with a situation can have comparable effects on adolescents and adults (Millstein and Halpern-Felsher, 2001). In an exhaustive review, Clifford (1991) relates risk-taking (measured with self-assessed questionnaires and price list designs) and educational choices, on children, teenagers and adults. It is shown that among other predictors such as motivation, game contexts or payoffs framing, risk-taking tasks have the advantage of facilitating learning and in-

crease children's level of academic achievement. She also explains that their tolerance for failure can be increased as their opportunities for taking risk increase as well. By comparison, conclusion drawn from adults are very similar. Hence, the difference between adults and teenagers might rather be on the evaluation of a consequence related to a risky situation and not on the perception of the risk itself.

This paper focuses on adolescents' risk preferences yielding first the question of how to properly measure them. The procedure created by Holt & Laury (2002) is widely used in experimental economics but had not yet been used on young individuals¹. Hence, we decided to implement this method on a sample of teenagers, first in order to verify if it can be applied to this specific sample, and second, to analyze determinants of risk aversion at this age. Risk preferences of young adults and even children can have a real impact on schooling decisions and future savings but also on risky behaviors such as smoking or drinking. Steinberg (2007) explains for instance that it would be more efficient to develop strategies such as raising the price of cigarettes, more vigilantly enforcing laws governing the sale of alcohol, or expanding adolescents' access to contraceptive services, in order to limit adolescent smoking, substance abuse or pregnancy, rather than policies aimed at changing adolescents' traits (short-sighted, impulsiveness...).

Most of the previous economic literature has studied adults' risk preferences and tries to explain it. Dohmen et al. (2010), for example, measure risk aversion in adults through the choice between a paid lottery and different safe payments. They also use two cognitive tests in order to link cognitive abilities to the willingness to take risks. Their main finding is that there exists a positive relationship between these two measures. In other words, individuals with lower cognitive abilities are significantly more risk averse.

However, recent literature tries to explain risk attitudes of children and their impact

1. This risk elicitation method has been mainly applied on undergraduate students and adults.

on educational outcomes or on risky behaviors. The purpose of these studies is not to evaluate how one can change the perception of risk by teenagers but to understand it, which can help at better communicating and elaborating policies for this specific population. Harbaugh, Krause and Vesterlund (2002) show how risk attitudes can change with age and more precisely that adults' and children's decisions significantly differ in gain or loss framework. They suggest that risk attitudes vary during a lifetime. On the other hand, Sutter et al. (2013) link lab measurements on ambiguity aversion, time preferences and field behaviors (such as savings, alcohol consumption, smoking etc...) in teenagers. They find a weak correlation between attitudes towards ambiguity measured in the lab and field behaviors, but a strong correlation with impatience.

Some others are more interested in the environmental impact on shaping risk preferences. Eckel et al. (2012) look at the relationship between risk aversion, school characteristics and cognitive skills measured by a math-related decision-making test. They base their study on 9th and 11th graders and find that school variables only impact risk preferences of the 11th graders.

The main question we try to answer in this paper is whether children's individual characteristics and specific cognitive abilities (usually difficult to measure) can explain risk aversion. Risky choices can be linked to pure cognitive abilities as mathematics or language knowledge. Moreover, we also think that creativity may decrease risk aversion as creativity drives innovation, which often carries risk. In this case, the pattern would be that creative children become innovative adults. How then is this related to risk preferences? This question is highly studied in management and industrial economics where innovation is crucial (Sternberg and Lubart, 1993, Heunks, 1998).

We elicit risk preferences with the Holt & Laury lottery choices (2002) where subjects have to choose between a safe lottery (safe choice) and a risky lottery (risky choice). The choice is always made between the same lotteries but the probabilities for both gambles are equal and vary by steps of 0.10 to 1.00. We apply this measure on 99

9th grade pupils coming from different classes of one middle school². To our knowledge, it is the first time these types of lottery choices are given to teenagers. For simplicity, we multiply the payment of the Holt & Laury lotteries by 10 so that payments become more realistic for children and represent higher stakes, however this might increase the degree of risk aversion³. This procedure allows us to obtain individual risk aversion that we try to explain with different creativity variables and personality traits.

The number of safe choices indicates the degree of risk aversion and the presentation of the lotteries are made such that consistent subjects should switch only once from the safe lottery to the risky lottery. However, as there are no forced choices here, this measure can yield "inconsistency" if the subject switches more than once implying a violation of expected utility but also if she chooses a lower sure payoff (a detailed description of the inconsistency measure is made later in the paper). Hence, another goal of this paper is to verify the amount of inconsistent subjects and to explain this inconsistency. One may be sceptical as to teenagers' ability to understand the lottery choices, however it has been shown that even children are able to understand probabilities (see [Reyna and Brainerd \(1995\)](#) for a review).

After having answered the lottery choices, pupils are asked to answer a personality questionnaire: the Brief Big Five ([Barbot, 2012](#)). It is a 100-item questionnaire that measures the 5 personality traits of the Big Five questionnaire (agreeableness, conscientiousness, extraversion, emotional stability and openness) but that is specifically designed for children and teenagers.

During two other in-class sessions, creativity scores are obtained using the Evaluation of Potential Creativity (EPoC) procedure ([Lubart, Besancon and Barbot, 2011](#)), a four-dimension procedure that is different from traditional one-dimension procedures

2. In France, 9th grade is the last year of middle school, just before entering high school. Pupils are 14 or 15 years old on average.

3. [Holt and Laury \(2002\)](#) use different payment levels to see if there is any effect of the magnitude of payoffs on risk aversion. They find that the level of risk aversion increases as the level of payoffs becomes higher.

such as the Torrance Test of Creative Thinking. Finally, we were able to gather the pupils' schooling grades during 8th and 9th grade in every subject. We prefer to use real schooling grades as opposed to IQ or cognitive tests (used in other studies) because we consider them to be more indicative of their real level of knowledge. Real grades also generate the beliefs that each pupil has of their relative level in the class. Plus, these are the main determinants for their future choice opportunities which can also rely on risk aversion. Schooling grades, especially in 9th grade, are going to be used for future educational choices when records will be examined by high school's principals and teachers. Hence, risk aversion or other kinds of preferences might also be explained by this type of school-specific abilities.

Our main finding is that the pattern of risk aversion using the H&L procedure is not very different from adults found in other papers. We observe that girls are more risk-averse than boys⁴. This result is mainly driven by two personality traits: girls who have a higher openness score and a lower emotional stability score tend to be less risk averse compared to boys. One of the creativity dimensions measure (showing the ability of producing many different original and appropriate ideas) also decreases the propensity of being risk averse, which is in line with the idea that being more creative leads to taking more risks. Concerning inconsistency, the rate of inconsistent subjects fits previous studies (26.26%)⁵ therefore showing that teenagers are not really more inconsistent than adults. Scientific grades are negatively associated with the probability of being inconsistent whereas grades in the humanities are positively associated. In addition, a higher score in two dimensions of the creativity measure based on verbal tasks decreases the probability of being inconsistent.

4. This is a traditional result from experimental economics looking at gender differences in risk preferences. Croson and Gneezy (2009) and Eckel and Grossman (2008) review this literature showing that in most of the studies (in economics), women are found to be more risk averse than men. Clifford et al. (1991) find that girls take lower academic risks.

5. Lévy-Garboua et al. (2012) find that 23.6% of subjects are inconsistent while doing the same framing H&L elicitation. Ballinger and Wilcox (1997) report a median switching rate of 20.8%, Loomes and Sugden (1998) report an average inconsistency rate of 18.3%, Blavatsky (2010) finds that overall 21.4% made inconsistent choices.

The paper is structured as follows: section 2 carefully describes the data (risk aversion measures, the BB5 questionnaire, the creativity test and the schooling grades), section 3 displays the results on risk preferences, section 4 presents the results on inconsistency and section 5 discusses the results and concludes.

2. The data

Data was collected on six 9th grade classes from a middle school in Rosny-sous-Bois, at the beginning of the year 2012. It is situated in the Parisian suburb and most of the pupils belong to low socio-economical status⁶. The results found in this study may be specific to this sample and would greatly benefit from replication in other types of schools.

In this section we review the four types of measures made with our sample of teenagers. We were not able to incentivize any task as the middle school did not allow us to give anything directly to the pupils. However, we told each class that if they were to go through all the sessions without any trouble and carefully, we would give three €15 gift cards to the head teacher that will be used to buy something for the class. Holt and Laury (2002) implement a treatment where they use hypothetical payoffs and compare these choices to real incentive payoffs. They do not find significant differences between hypothetical and real incentivized choices so we can assume here that hypothetical payoffs will not bias our sample's choices. We went into each of the six classes, for 50 minutes, three times: twice for the creativity test and once for the risk aversion measure. We also came back few months later to present a brief summary of our research.

6. We asked pupils the socio-professional category of their parents as well as their exact job. We were then able to establish the proportion of high qualified mothers and fathers in our sample which is respectively 35.4% and 29.1%.

2.1. *Risk Aversion Measure*

99 pupils were asked to answer ten paired-lottery choices, based on the [Holt and Laury \(2002\)](#) procedure. Thus, the choice was between a safe lottery and a risky lottery. The original lottery choice of Holt & Laury is between the safe lottery of \$2.00 and \$1.60 vs. the risky lottery \$3.85 and \$0.10, while varying the probabilities for each choice. In this experiment, payoffs are multiplied by 10 and choices were displayed simultaneously on a computer. As mentioned in the introduction, multiplying the payoffs by 10 can increase risk aversion and can refute constant relative risk aversion (see [Holt and Laury \(2002\)](#)). However, it is not an issue here as we apply the same measure to all of our sample so they all face the same framing and our comparisons are made on characteristics differentials and not on the perception of payoffs (hypothetical vs. real, high vs. low...).

The experiment was conducted in the classroom, with the pupils being in front of a notebook computer where lottery choices were displayed. The instructions for lottery choices were read aloud and the pupils could also read them on the screen. It was presented as a choice between two scratch cards with different chances of winning. The instructions were the following: "We consider two scratch games with different outcomes. The chances of winning each outcome vary. There are 10 scratch games to be played. With the first scratch game you can either win €20 or €16. In a second scratch game you can win either €38.50 or €1.00. You will have then to choose between these two scratch games 10 times (game A or game B)". Then, an example with different payoffs was given to make sure they understood the task and they were allowed to ask questions. [Figure 4.1](#) shows the screen pupils saw with the 10 different "games". However, even though the sure choice of winning €38.5 must be preferred to the sure choice of winning €20 (corresponding to the 10th choice), no forced end point nor any forced choices after the first switch were implemented to allow for inconsistency.



Figure 4.1: Lottery choices

2.2. The BB5 questionnaire and demographics

After the lottery choices, teenagers are asked to fill in a demographic questionnaire (gender, age, parental educational background, parents' occupation) and the Brief Big Five (BB5).

This part of the experiment focuses on personality traits. Personality can be defined as "what defines what a person will do when faced with a defined situation" (Cattell, 1979). The Five-Factor Model (FFM) of personality traits (Big Five) is certainly the dominant theory in personality research (McCrae and Costa, 1987, John and Srivastava, 1999). The starting point of this theory is natural language; more precisely, the way people describe their personality or others' personality (Saucier and Goldberg, 1998). Personality vocabulary provides a finite set of attributes for personality description (Goldberg, 1981). The Big Five model was discovered in analysis of several

languages (Saucier and Goldberg, 1998) and its usefulness utility and accuracy have been confirmed by previous literature. It proposes a structure of human personality in five dimensions:

- Agreeableness: Altruism, sympathy, cooperative, helpful.
- Conscientiousness: Organized, achievement-oriented, disciplined.
- Extraversion (vs. neurotism): higher degree of sociability, emotional activity seeking, talkative.
- Emotional Stability (vs. neurotims): impulse control, not anxious, calm.
- Openness: Intellectual curiosity, preferences for novelty and variety.

As our subjects are teenagers, we use the Big Five inventory targeted for teenagers called *The Brief Big Five* (BB5) (Barbot, 2012). This questionnaire allows us to get the score for these same dimensions. The BB5 consists of a list of 100 adjectives and highlights the personality traits that characterize an individual. For each adjective, they have to declare on a 5 point-Likert scale if it "totally" to "not at all" corresponds to them.

Among the 99 pupils who filled in the BB5, three did not finish the questionnaire⁷, and five did not have valid scores⁸, we thus have 91 observations for the BB5.

To facilitate interpreting the effect of personality traits on our different measures, we use *Standard Scores*. We follow a specific standardization of the row score in order to obtain a score scaled on 11 classes for each of the five personality traits. This 11 class categorization is precise enough to allow us to compare individuals and to obtain an identical scale for each trait.

7. This means that they did not do the final click which would have recorded all their answers into the computer. Hence, we do not have any answers for these subjects as none of them have been recorded.

8. In addition to the five traits, there are two scores that identify bias in responses: the acquiescence tendency (too many extreme answers) and neutral responses (no positioning).

As we will be interested in gender difference towards risk aversion, table 4.1 reports the BB5 scores (based on standard scores) by gender and the p-values of a two-tailed t-test. This table shows that girls are significantly more agreeable, more open and slightly more conscientious than boys. There are no significant differences for extraversion and emotional stability between genders⁹.

	Girls	Boys	T-test p-value
Agreeableness	5.62 (2.26)	4.33 (2.27)	0.00
Conscientiousness	6.64 (2.17)	5.88 (1.95)	0.08
Extraversion	5.78 (2.06)	5.37 (1.76)	0.30
Emotional Stability	6.55 (2.62)	6.53 (2.25)	0.97
Openness	6.00 (2.00)	5.00 (2.30)	0.03
Observations N	42	49	

Table 4.1: BB5 standard scores by gender (standard deviation in parentheses)

2.3. Measuring creativity via the EPoC procedure

EPoC (Lubart, Besancon and Barbot, 2011) is a procedure that measures the creative potential of pupils from elementary to middle school. It is a multivariate approach that uses verbal and graphic tasks measuring two types of creative thinkings: *Divergent Thinking (DT)* and *Integrative Thinking (IT)*¹⁰.

Divergent thinking is a thinking process that consists of generating the maximum number of creative solutions. The EPoC procedure embodies two types of DT:

- Graphical Divergent Thinking (*GDT*): pupils are asked to produce a maximum

9. Previous studies made on adults found that women were more agreeable and more conscientious, and men more emotionally stable. No constant results are really found for openness and extraversion (see Costa Jr, Terracciano and McCrae (2001), Rubinstein (2005), Barbot (2012)).

10. A precise description of the tasks can be found in the appendix of part 2 of the dissertation.

quantity of original drawings in 10 minutes based on a simple shape.

- Verbal Divergent Thinking (*VDT*): the experimentalist gives a beginning or an end to a story and pupils are asked to produce a maximum quantity of story endings or beginnings in 10 minutes.

Integrative thinking is a cognitive activity that consists of combining many elements. It is a process of idea association as well as a selective comparison and combination process that allows for the synthesis and convergence of a unique and original production. Two types of IT are thus distinguished:

- Graphical Integrative Thinking (*GIT*): From ten drawings the pupil is asked to produce one unique drawing using at least three drawings from the list in 10 minutes.
- Verbal Integrative Thinking (*VIT*): The experimentalist gives a story title and the pupils have 10 minutes to finish the story in the most original manner.

Creativity is measured during two sessions in which we implement two versions of the test for robustness. The tasks are the same but the content differs (different drawings, different stories...). During each session, both types of thinking are measured. One session lasts 50 minutes on average and was done in the classroom with paper, pencils and felt pens (see figure 4.4, 4.5, 4.6 in the appendix for different outputs of VIT, VDT and GIT). In summary, in the DT tasks, pupils are asked to produce a maximum of ideas in response to a unique stimulus. In the IT tasks, they have to produce a more elaborated synthetic solution.

Creativity scores for each type are based on the number of outputs (for DT tasks) and the degree of originality. The higher the score, the higher the individual's creative potential. For an easier analysis we use standardized scores of these four measures.

2.4. Grades collected

In order to have a measure of schooling achievement, we obtained average grades of 90 pupils for all subjects and each trimester of their 8th and 9th grades (or 6 trimesters)¹¹. The average grades we obtained are on a 20 points basis (usual French grading). For our estimations we use the standardized average grades based on the six trimesters¹². For our analysis on inconsistency we use the average grades in scientific subjects (math, physics and chemistry, biology) and in humanities (French, history and geography, world language). This aims to separate specific domains where computations skills are needed on one hand (in science) and less useful (in humanities). Overall, girls have higher grades than boys. Figure 4.2 shows the distribution of scientific and humanities scores by gender.

3. Attitude toward risk

Along this section we analyze risk aversion on the whole sample of subjects but also on the ones considered to be consistent, that is to say subjects who do not make inconsistent choices (see section 4 for a detailed explanation). Moreover, as we will use creativity scores and BB5 scores in our estimations, our sample is thus reduced to pupils who responded to the lottery choices as well as the ones who did the creativity test. This leaves us with 72 valid observations when not taking into account inconsistency, and 52 valid observations for consistent subjects in order to estimate risk aversion. This may raise a problem of selection bias: pupils who attended the three sessions may have different characteristics than pupils who only attended one or two sessions

11. Note that for each of the measures we do not have the same number observations. Within these observations, we do not always have all the observations for each measure and for each pupil. That is why the sample size varies according to the measure considered.

12. The evolution of average grades in each subject does not significantly vary across time. So we can use the average grade on the six trimesters.

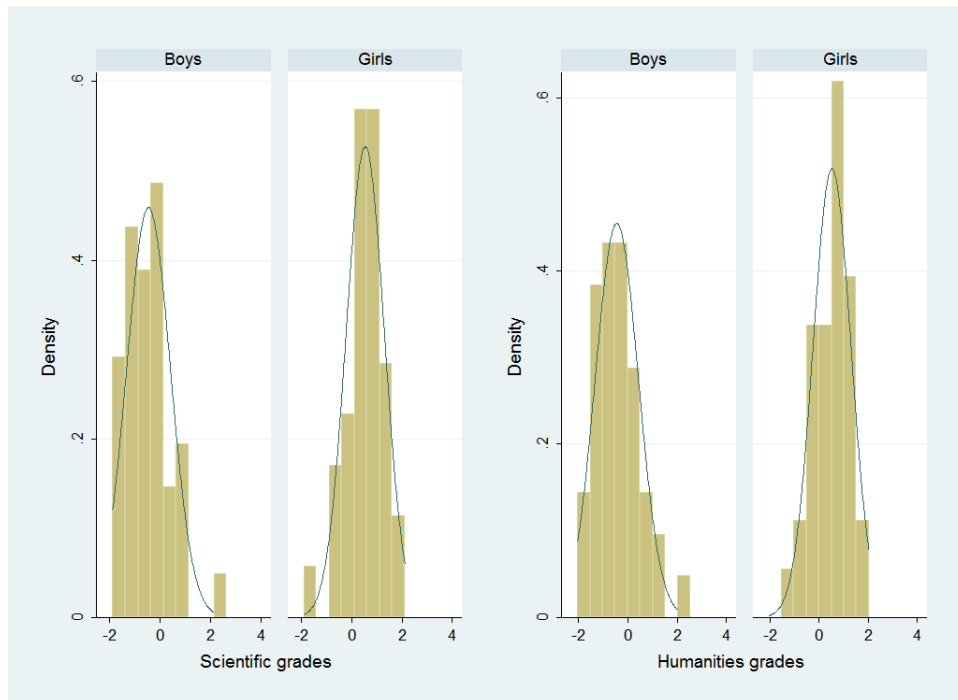


Figure 4.2: Distribution of scientific and humanities scores, by gender (n=90)

(more conscientious¹³, more diligent). The analysis is thus run only on the pupils who attended the three sessions.

We use Holt & Laury's procedure for defining at which decision subjects switch from the safe lottery (game A) to the risky lottery (game B). Figure 4.3 displays the proportion of safe A options chosen by all the pupils (solid line) and the proportion of safe A options under the assumption of risk neutrality (dashed line), as well as broken by gender. The shape of the curve is consistent with previous studies that use adults as respondents.

We find that our sample of teenagers globally exhibit risk aversion as, after pair 4, they more often choose the safe option than if they were risk-neutral. We can also see that girls and boys are more risk averse, which is not just driven by girls, but that the latter have a higher proportion of choosing the safe lottery until pair 7 (the blue line

¹³. No significant difference was however found in the conscientiousness score between those who attended the three sessions and those who did not

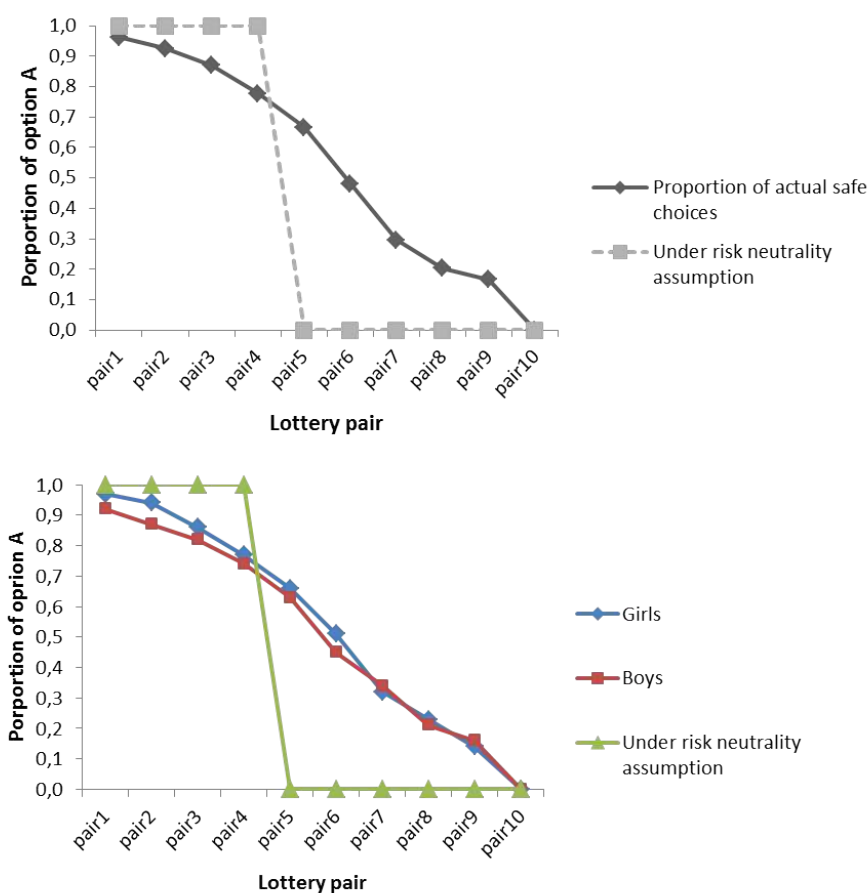


Figure 4.3: Proportion of safe choices in each decision

stays above the red line until pair 7).

We use a functional form of a Constant Relative Risk Aversion (CRRA) utility function that allows us to obtain the risk-aversion parameter r such that $U(x) = \frac{x^{1-r}}{1-r}$ if $r \neq 1$ and $U(x) = \log(x)$ if $r = 1$. This specification is usually chosen in papers analysing risk aversion based on this type of task. We are then more able to compare our results to theirs by using it as well. The coefficient of risk aversion is r , and x is the payoff. Table 4.2 reports the different values of r according to the number of safe choices (number of times game A is chosen), and the proportion of subjects for each of these choices for the whole sample and broken by gender. r takes negative values for

risk lovers, positive values for the risk-averse and is equal to zero for risk neutrality. We include in this table only the pupils who exhibit complete consistency in answering the lottery choices. We can see that a higher proportion of boys exhibits a lower number of safe choices. However, a Fisher exact test yields a p-value=0.954 meaning that there is not a statistically significant relationship between the distribution of the number of safe choices and gender.

Number of safe choices	Range of Relative Risk Aversion (r)	All	Girls	Boys
0-1	$r < -0.95$	0.11	0.07	0.15
2	$-0.95 < r < -0.49$	0.05	0.07	0.03
3	$-0.49 < r < -0.14$	0.03	0.20	0.17
4	$-0.14 < r < 0.15$	0.19	0.21	0.17
5	$0.15 < r < 0.41$	0.17	0.14	0.20
6	$0.41 < r < 0.68$	0.16	0.21	0.12
7	$0.68 < r < 0.97$	0.11	0.14	0.09
8	$0.97 < r < 1.37$	0.05	0.03	0.06
9-10	$1.37 < r$	0.13	0.10	0.15

Table 4.2: Risk aversion and proportion of safe choices

We now try to explain risk aversion with personalty traits and creativity scores. With this we consider the question: are more creative students prone to risk taking? Moreover, since it has been found that women are more risk averse than men we ask: is this the case among young individuals?

We use the interval regression model (Coller and Williams, 1999, Harrison and Rutström, 2008) based on the different risk aversion coefficients intervals. The dependent variable is the CRRA interval that each pupil implicitly chooses when switching from the safe choice A to the risky choice B (see table 4.2). Using this model, we control for individual characteristics such as gender and personality traits. Table 4.3 presents estimates of this interval regression model in the whole sample including inconsistent subjects, and on the restrictive sample of only consistent subjects¹⁴. The first two

14. OLS regressions were run on the number of safe choices and on the switch time as well as interval regression on the probability intervals. They yield the same results.

columns presents the estimates when controlling for gender and age. On the whole sample, the age coefficient is positive and significant at a 5% level, meaning that the older, the more likely to be risk averse¹⁵. In the second and third column we add controls for personality traits and creativity scores. The estimates for girls and age are positive and significant at 10%. Girls are therefore slightly more risk averse than boys as well as older pupils (a higher CRRA coefficient implies greater risk aversion). We find no direct effect of personality traits nor creativity on risk preferences.

Understanding why women are more risk averse than men is still a complex investigation. In order to determine if this effect is driven by personality we add interaction terms between each personality traits and the female dummy variable. We find that girls who exhibit more emotional stability than boys will be more risk averse and this is robust on both the whole sample and subsample (columns (5) and (6)). On the less restricted sample (column (5)) the estimate of the interaction between gender and agreeableness is negative and significant at a 10% level implying that a higher score of agreeableness yields a lower aversion to risk for girls, compared to boys. However, on the more restrictive sample (column (6)) this estimate becomes insignificant but the positive and significant estimate of the interaction term between gender and openness imply that girls who are more open, compared to boys, will be less risk averse¹⁶. This suggests that girls are indeed more risk averse than boys but that personality can mitigate this result; the differences in personality traits between boys and girls can influence risk aversion. Moreover, as results on openness and agreeableness change according to the sample size, this may indicate a difference in personality between the whole sample of pupils and the consistent pupils. However, a two-tailed t-test yield a p-value > 0.1 when testing for differences in scores of openness and agreeableness between the types

15. Age is in years. The mean age is 14.4. We did not have access to their exact date of birth and the sessions when we asked for their age was done in January 24th and 25th. The distribution of age is the following: 2.7% of our sample is 13, 64.0% is 14, 28.0% is 15 and 5.3% is 16.

16. [Borghans et al. \(2009\)](#) study risk aversion and ambiguity aversion on 16 year olds. They find an effect of these same personality traits on risk aversion, but when separating their sample by gender this effect disappears.

VARIABLES	All sample (1)	Consistent (2)	All sample (3)	Consistent (4)	All sample (5)	Consistent (6)
Girls	0.249 (0.210)	0.273 (0.226)	0.408* (0.232)	0.479* (0.282)	0.622 (0.853)	-0.083 (0.727)
Age	0.392** (0.173)	0.277 (0.186)	0.232 (0.176)	0.304* (0.177)	0.090 (0.182)	0.107 (0.188)
Agreeableness			-0.001 (0.041)	-0.034 (0.035)	0.067 (0.064)	0.034 (0.069)
Conscientiousness			-0.019 (0.052)	-0.035 (0.052)	-0.021 (0.080)	-0.100 (0.084)
Extraversion			0.004 (0.062)	-0.021 (0.052)	0.078 (0.096)	0.093 (0.116)
Emotional Stability			-0.005 (0.056)	0.064 (0.047)	-0.176** (0.069)	-0.169*** (0.064)
Openness			-0.016 (0.058)	-0.006 (0.062)	0.044 (0.069)	0.094 (0.067)
GIT			-0.040 (0.065)	-0.003 (0.056)	-0.052 (0.062)	0.007 (0.056)
VIT			-0.035 (0.083)	0.060 (0.109)	-0.028 (0.078)	0.061 (0.085)
GDT			-0.021 (0.076)	-0.108 (0.085)	-0.116 (0.096)	-0.252*** (0.098)
VDT			-0.094* (0.050)	-0.070 (0.047)	-0.082 (0.051)	-0.054 (0.046)
Girl*A					-0.182* (0.095)	-0.147 (0.099)
Girl*C					-0.054 (0.100)	0.044 (0.105)
Girl*E					-0.131 (0.123)	-0.107 (0.137)
Girl*SE					0.381*** (0.105)	0.410*** (0.099)
Girl*O					-0.128 (0.097)	-0.195** (0.090)
Constant	-4.926* (2.515)	-3.573 (2.756)	-2.498 (2.761)	-3.983 (2.865)	-0.327 (2.909)	-0.651 (3.089)
lnsigma	-0.142 (0.114)	-3.17*** (0.148)	-0.190 (0.121)	-0.364** (0.153)	-0.258** (0.121)	-0.481*** (0.159)
Observations	75	52	72	52	72	52

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4.3: Interval regression on the CRRA coefficient

of sample.

When looking at creativity scores, when including them in column (3), verbal divergent thinking (VDT) negatively predicts risk aversion (the estimate is 10% significant) meaning that a higher score in this dimension decreases the propensity of being risk averse. The other creativity scores' estimates are negative as well, but non significant. When adding personality and gender interaction terms, we find that a higher score of graphic divergent thinking (GDT) decreases risk aversion but only on the sample of consistent subjects. Results on creativity potential are less consistent across the different samples but the negative signs of the estimates matches the idea that creative people are less risk averse.

4. Explaining inconsistency

In addition to explaining adolescents' risk aversion we try to understand why some among them exhibit inconsistency in lottery choices. [Lévy-Garboua et al. \(2012\)](#) show that the framing effect influences inconsistency. For example, they find that simultaneously presenting the lottery choices significantly decrease inconsistency compared to a sequential framing. In this section, we are more interested in pointing out which individual characteristics can increase or attenuate inconsistency. We categorize subjects as inconsistent on the lottery choice task based on two criteria:

- The number of switches¹⁷: a subject who makes two or more than three switches. We allow subjects who switch three times to be consistent and consider that their certainty equivalent is situated between the first and the third switch. By switching three times, a subject can show either an indifference between the first and the last switch row or a doubt on their preference implying this multiple switch. These subjects usually switch back quickly so we can neglect the between-choice.

17. By switching we mean the moment a subject switch from option A to option B.

- The last choice: as we did not oblige subjects to choose the option B (a sure payoff of €38.5), subjects who did not make this choice are also included in the inconsistent sample¹⁸. 10% of the sample actually chose the option A at the tenth choice.

Of the 99 pupils who answered the lottery task we obtain 26.26% of inconsistent subjects (26 subjects). As mentioned in section 1, this rate does not highly differ from the one found in previous papers where the decisions were made by adults¹⁹. Again, we might face a problem of selection bias. When taking into account participation in the creativity and BB5 sessions, the number of remaining observations decreases to 75. Thus we decided to run estimations on the pupils who came to all of the sessions, which concludes an inconsistency rate of 27.8% (21 subjects). A two-tailed proportion Z-test shows that these two rates are not significantly different ($p=0.4$).

To get a better understanding of how to explain inconsistency, we run a probit model on the variable of being inconsistent (=1 if the subject is inconsistent, 0 if she is not). We expect that cognitive abilities can be a real determinant on this probability. We hypothesize that scientific knowledge (math, physics, biology) can decrease the probability for pupils of being inconsistent, especially through the knowledge of probability computation. We hence use the mean standardized grade in scientific courses i.e. math, physics and biology. To balance this hypothesis, we also compute the mean standardized grade in the humanities: French, history & geography and world languages. These average grades are based on 8th and 9th grade from all 6 trimesters. We also use creativity scores, which are also a cognitive ability that could be linked to the probability of being inconsistent. Since integrative thinking is a measure of the ability to produce or synthesize in a creative manner and relies on logic and the ability to solve, we expect to find a negative association with the probability of being inconsistent.

18. Although they could be considered as subjects that simply did not understand the task.

19. See Ballinger and Wilcox (1997), Blavatsky (2010), Lévy-Garboua et al. (2012), Loomes and Sugden (1998)

Table 4.4 shows the probit estimation on being inconsistent. We add personality traits from the BB5 questionnaire as controls in addition to age and gender. However, as we assume inconsistency to be the consequence of a lack of cognitive ability, we expect no effect of personality on this variable.

We can see that in the first column of the probit table the estimate of scientific grades is negative and highly significant, implying that the higher the grade in these subjects the lower the probability of being inconsistent. Humanities' grades have no direct impact on the probability of being inconsistent. When adding the four creativity scores to the set of explanatory variables (column (2)), we find that having a higher score in the verbal integrative thinking decreases as well the probability of being inconsistent. The significance of the scientific grades' estimate is still high and thus robust.

In column (3) we add the five personality traits: no effect is found here. However, adding these controls implies a positive and significant association between grades in the humanities and the probability of being inconsistent. In other words, scientific knowledge can decrease the probability of being inconsistent on the H&L task, whereas better grades in humanity subjects can increase it. Concerning creativity, the VIT estimate remains negative and significant and the estimate of verbal divergent thinking being also negative, become slightly significant (at a 10% level). This indicates that verbal creativity might have a greater impact on reducing the probability of inconsistency compared to graphic creativity that does not have an effect here.

Cognitive abilities may reduce the probability of inconsistency, measured here by creativity and schooling grades, however, personality has no impact. Being inconsistent may be due to framing issues but also to individual characteristics and cognitive abilities. It would be therefore interesting to run the same analysis on adults or younger children.

VARIABLES	Estimate (1)	Estimate (2)	Estimate (3)
Girls	0.254 (0.396)	0.745* (0.441)	0.807* (0.467)
Age	0.124 (0.287)	-0.494 (0.343)	-0.734* (0.444)
Humanity subjects	0.545 (0.359)	0.613 (0.380)	0.745** (0.368)
Scientific subjects	-0.945** (0.386)	-1.035** (0.404)	-1.186*** (0.402)
GIT		-0.097 (0.137)	-0.172 (0.155)
VIT		-0.427*** (0.150)	-0.416*** (0.158)
GDT		0.117 (0.116)	0.107 (0.124)
VDT		-0.170 (0.108)	-0.198* (0.117)
Agreeableness			0.007 (0.097)
Conscientiousness			-0.117 (0.116)
Extraversion			-0.124 (0.099)
Emotional Stability			0.063 (0.111)
Openness			0.005 (0.105)
Constant	-2.579 (4.205)	6.011 (5.003)	10.304 (6.856)
Observations	70	70	70

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: All variables are continuous, except female which is binary and equal to 1 for girls, 0 for male. Observations are equal to 70 as we have lottery choices, creativity measures and grades for only 70 subjects out of the 99 that answered the H&L measure.

Table 4.4: Probit on being inconsistent

5. Discussion and conclusion

Risk aversion is a crucial topic in economics and has been widely studied on adults in experimental economics. There are different methods of risk preferences elicitation. Psychologists often use questionnaires to determine the propensity of risky behaviors in individuals. Rather than using self-declared questionnaires, economists prefer gambling tasks with money incentives. The Holt & Laury (2002) procedure is one of the main elicitation procedures used in experimental Economics. It consists of a series of choices (with increasing probabilities) to be made between a safe lottery and a risky lottery. It has been shown that this method is correlated with risky economic decisions. However, risk aversion is not only related to adulthood decisions but also on children's having then an impact on their own decisions, especially regarding schooling, future career decisions or any risky behaviors (smoking, drug consumption...). It seemed important to us to first investigate whether such risk elicitation procedure could be applied to teenagers, and second, to investigate what kind of individual variables might impact their risk preferences.

The psychology of creativity literature shows that creativity is highly associated with risk seeking. The role of creativity at schools may thus shape risk preferences. This is also the case for personality traits and cognitive abilities. Hence, we conducted studies in a school and we were able to obtain these different types of measures on 9th graders. We measured their creativity based on the EPoC procedure, and the Big Five questionnaire. We were therefore able to obtain five personality scores, gain access to their reports cards, in order to have a measure of their cognitive abilities. Finally we implemented the H&L lottery choices.

Our paper shows that it is possible to use this procedure on teenagers. They are able to understand it and do not yield a higher inconsistency rate than adults. The way the test is presented to them and the vocabulary used are important and must be

carefully thought. We find that girls are slightly more risk averse than boys, which is in line with previous findings. However, it seems that this finding is modulated through some of the pupils' personality traits: if girls have a higher openness score they are less risk averse than boys. Girls who present a higher score of emotional stability are more risk averse than boys. Personality has an effect, but only through interactions with gender. This implies that risk aversion that is generally more observed among girls may be decreased by influencing personality traits that promotes risk taking, such as openness.

As expected, a higher level of creativity decreases the level of risk aversion, but this is only found to be true for one creativity subtype, the graphical divergent thinking. Increasing creativity in children can also modify risk preferences and have a future impact on schooling choices and later on the labor market. However, another study based on the same dataset (Berlin, Besançon and Tavani, 2013), shows that verbal divergent thinking is negatively associated with most of the schooling subjects. This raises an important question as to the role educational system chooses to give to creativity at school and how it can be combined with the good level of knowledge in fundamental subjects in order to create optimal decision makers.

Another aim of this paper is to explain inconsistency. Indeed, contrary to other elicitation procedures, the H&L lottery choices imply a certain rate of inconsistency observed in every study using this method, when no forced choices are implemented. Teenagers may be expected to be more inconsistent than adults. Nevertheless, the inconsistency rate we find in this paper is not much higher than those found for adults in other studies. In order to refrain the amount of inconsistency on this type of task, it is however advised to force choices to only one switch, so measure of risk preferences can become more precise. We are interested in understanding which individual characteristics can decrease the probability of being inconsistent. This paper does not take into account the framing effect but shows that individual characteristics can also affect the

probability of being inconsistent. This is the case for grades in scientific and humanity subjects (respectively decreasing and increasing the probability), but also the ability to synthesize into a creative manner through verbal divergent thinking. The opposite sign of the estimate found for scientific and humanities grades show that math skills, or at least computation skills, might be needed to reduce inconsistency. This can also raise the question of how individuals, whatever their age, perceive probabilities conditional on their level of the domain-specific cognitive skills. No effect of personality was found there. Nevertheless, this result may be different when looking at adults which can be a further work to investigate.

Further studies could be done on a larger sample of teenagers, from different school and age. This would inform on the temporal variation of these results. Are they robust on younger children? Does using real incentives or modifying the payoffs change the results? These are further studies which can be explored.

D. Appendix



Figure 4.4: Two examples of GIT task.

La goutte d'eau.

Kaylinn regardait tranquillement une émission à la télé. Tout d'un coup, son frère Kevin lui lance de l'eau sur le visage. alors, Kaylinn se lève, puis lui court après dans toute la maison, Kevin se cache, va dans la cuisine et prend un oeuf. Bien caché, il attend d'avoir Kaylinn dans sa ligne de tir. Quand il la voit enfin, il lui jette l'oeuf dessus. S'en est trop, "C'est la goutte d'eau qui fait déborder le vase," dit elle. Elle attend alors que Kevin sorte de sa cachette, mais en vain. Le lendemain matin, la jeune fille voit Kevin dans son lit. Elle va chercher de l'eau, puis à son tour lui jette de l'eau. Kevin sursaute, et tous les deux se mettent à rire.

La goutte d'eau.

C'était un jour de pluie, et je voulais aller dehors pour m'amuser dans boue, mais maman ne voulait pas. Mais moi je suis têtue, alors je me précipite dehors pour m'amuser avec mes copains. Plus le temps ^{passé} plus je m'amuse, mais la pluie ne cessait de tomber. Au bout d'un moment la pluie était tellement forte qu'elle martelait le ~~de~~ ^{dessus} de la terre. Quand tous mes copains étaient rentrés je me retrouvais seul dans la boue sous une pluie battante. Quand tout à coup j'entendis un énorme bruit, je lève la tête lorsque j'aperçois une énorme goutte, plus grosse que moi qui arrivait à toute vitesse pour me percuter... et là, je me réveille en sursaut dans mon lit trempé de sueur. Ouff! - ce n'était qu'un rêve! À l'avenir j'écouterais maman!

Fin

Figure 4.5: Two examples of VIT task: writing a story based on the title "Drop of Water"

- Claude vit un éléphant marcher et écraser tout ce qu'il y avait sur son passage, d'un seul coup, il arrêta de jouer et rentra chez lui.
- Claude vit un immeuble s'effondrer
- Claude vit que ce bruit n'était que celui d'une jeune fille ne sachant pas chanter, puis il se moqua d'elle.
- Claude vit un tracteur rouler près du parc.
- Il tanna sa tête, et vit le sol fissuré
- Claude n'eut pas le temps de tourner la tête, qu'il était poursuivi par un chameau
- Claude rentra chez lui
- Mais il continua à jouer sans s'en occuper
- Il vit qu'un enfant avait cassé une des installations de l'air de jeux.
- Il se retourna, et vit, à terre, un chat qui lui semblait bien mal en point, il avait une voix bizarre.

- Un après Claude jouait au Parc lorsque il entend un bruit..
- ① Il se précipita apperça derrière un buisson, il n'osa même ^{pas} ouvrir les yeux. Claude prit son courage à deux mains puis sorti de sa cachette. Lorsque il apperçut que ce n'était petit chaton, il se senti ridicule.
 - ② Claude très curieux s'empêssa d'aller voir ce qu'il se passait, mais le bruit se déplaçait de plus en plus vite. Claude le suivait toujours sans savoir vraiment où il allait. Au bout d'un moment il arriva dans les bois et commença à avoir peur, quand tout à coup 3 jeunes ~~un~~ gars sortis des buisson. Mais ce n'était rien, juste ses amis qui voulait lui faire une peur.
 - ③

Figure 4.6: Two examples of VDT task: writing as many ends of story as possible

Appendix of Part II

Steps for the EPoC procedure

There are in total eight steps and two "forms". Two sessions of 45 minutes, that each gathers 4 steps. For both sessions, the tasks are the same but the stimuli are

different.

- i. Graphical Divergent Thinking task: Finding as many ends of story as one can, in 10 minutes. The pupil is asked to create as many drawings as he can from a specific shape. One drawing for one sheet of papers, but the pupil can use as many sheets as wanted. At the end of the task, the pupil has to give a title to each of his drawing.
- ii. Verbal Divergent Thinking task: The pupil is asked to invent as many ends of story as he can from a specific given beginning, in 10 minutes. In a second session, he is asked to tell story beginnings based on a story end. The experimenter points out that the ends or beginning of story have to be as interesting and original as possible as well as different from the ones the others children could write.

Examples:

- The experimenter reads a beginning of story about a child named Dominique who is with his/her grand-mother. The weather is nice and the grand-mother says to Dominique that he/she has to go to the grocery, leaving Dominique alone in the garden...
 - The end of the story given was: "And the apple fell from the tree." or "And the stone rolled far away."
- iii. Graphical Integrative Thinking task: the pupil is asked to make an original drawing in 15 minutes, different from what the others would do, using at least 4 shapes from different the abstract stimuli.
 - iv. Verbal Integrative Thinking task: The pupil has to invent a story from this title: "The Keyhole", in 10 minutes. Another version would be to tell a story in 10 minutes that includes a child, a fish and an unknown person.

The following documents are different outputs from two different children with different levels of creativity. This gives an idea of the different production these tasks could yield. They do not include Graphic Divergent thinking as they were done on very small sheets of papers, with pencils, so scanning was a hard task and not very visible.

* Le vent a soufflé sur le pommier,
l'un à l'autre les hommes tombèrent puis
[...]

* Un éléphant, qui avait faim, secourut le
pommier jusqu'à que [...]

* Le printemps se termina et l'été commença
jusqu'à que [...]

* Le sage donna une seconde chance à Chang
mais qu'il devait faire c'est prouver avant
que [...]

*



* Il était une fois un enfant et sa grand-mère qui vivait dans une petite maison.

Un jour l'enfant demande à sa grand-mère une histoire pour s'endormir. La grand-mère

lui raconte l'histoire avec l'oiseau.

Elle raconte que l'oiseau était un oiseau timide mais qui avait beaucoup

de courage. Un jour il y avait un autre oiseau fille qui était menacé

par un ~~faucou~~ faucon. L'oiseau,

qui depuis longtemps est amoureux

de cet oiseau, vole à son secours et

arrive à éloigner le faucon de

l'oiseau fille. L'oiseau fille remarque

le courage, tombe immédiatement

amoureuse de l'oiseau garçon.

Le petit enfant qui fut tout content

que l'oiseau a pu aimer l'oiseau

fille, s'endormit dans un profond

sommeil où il rêva que lui aussi

était un train de sauver une fille.

* Pendant que sa grand-mère était sortie, Dominique en profita pour aller prendre des gâteaux dans le placard. Mais comme Dominique était trop petit, il prit une chaise et monta pour atteindre le placard. Mais il perdit l'équilibre et fit tomber le placard. La cuisine était dans une état catastrophique. Mais ~~mais~~ Dominique n'eut pas le temps de se relever que sa grand-mère et lui le désordre. Dominique fit peur et dut ranger toute la cuisine.

78

* Dominique ne fut pas rassuré du fait que sa grand-mère était partie et qu'il était tout seul. Il s'imagina ~~que~~ que les kidnappers allaient le kidnapper et demander un rançon. Après cette réflexion Dominique alla se cacher dans une armoire. Quand sa grand-mère fut rentrée, il lui sauta dans les bras et l'aida de sauter. ~~La grand-mère~~

79

* Dominique qui n'aurait pas été seul, eut ~~la~~ sa grand-mère qui s'arrêta. Dominique rejoignit sa grand-mère et lui demanda si elle pouvait l'accompagner et celle-ci dit oui.

80

* Dominique attendit sa grand-mère en regardant la télé et des qu'elle

81

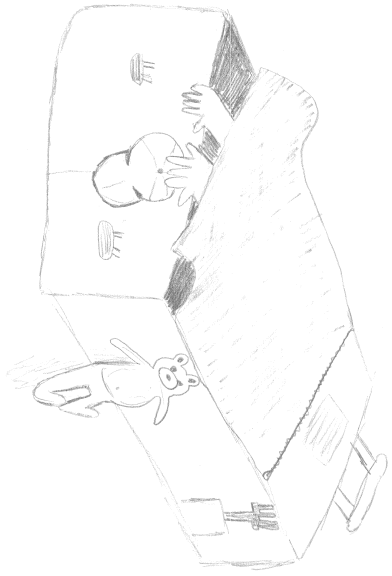


Annoy . A .
Sic

de trou de la serrure.

Rachid, un enfant de 11 ans, est un enfant très naïf. Il croit à tous ceux qu'on lui a dit depuis qu'il est tout petit. Il croit au Père Noël, au Sorcier, au monstre, au vampire, ... Mais il aime aussi raconter des histoires mais bien sur tout le monde ne l'écoute pas. Un jour, pendant qu'il monte l'escalier de son immeuble, il entend un bruit étrange dans l'étage où il se trouvait. Bien sûr, ce petit curieux veut savoir qu'est-ce que c'est. Et il marche vers la direction de ce bruit et s'arrête devant une porte. Mais la porte est fermée. Rachid cherche une solution et ~~trouve~~ voit que le trou de la serrure est suffisamment grande pour qu'il puisse voir de l'autre côté, lorsqu'il met son œil regardé, il aperçoit un monstre. Un monstre tellement hideux qu'il passe un cri effroyable de monstre voyant qu'il est observé cour vers la porte et capture Rachid. Rachid, effrayé, ne

- 1) Alors que Pierre donna un coup de pied dans la pierre...
- 2) Alors que des Racailles se jetaient des pierres...
- 3)



Un enfant, en bord de mer, alla se baigner sans l'aide de ses parents, il nagea loin, très loin de ses parents, jusqu'à ne plus les apercevoir.

Mais tout-à-coup, il sentit comme un élement qui l'emporta soudainement dans un étrange monde sous l'eau.

C'était une base géante, il n'y avait pas grand monde. Cet enfant, alla alors vers une large arche où il vit le Grand poisson géant. Le poisson était sûrement le patron des lieux, il avait un air grossier, et peu respectueux.

Mais vient l'arrivée de 2 requins marchant vers le petit garçon, ils le prit sur leurs épaules, et l'accompagna brusquement vers les cachots des prisonniers.

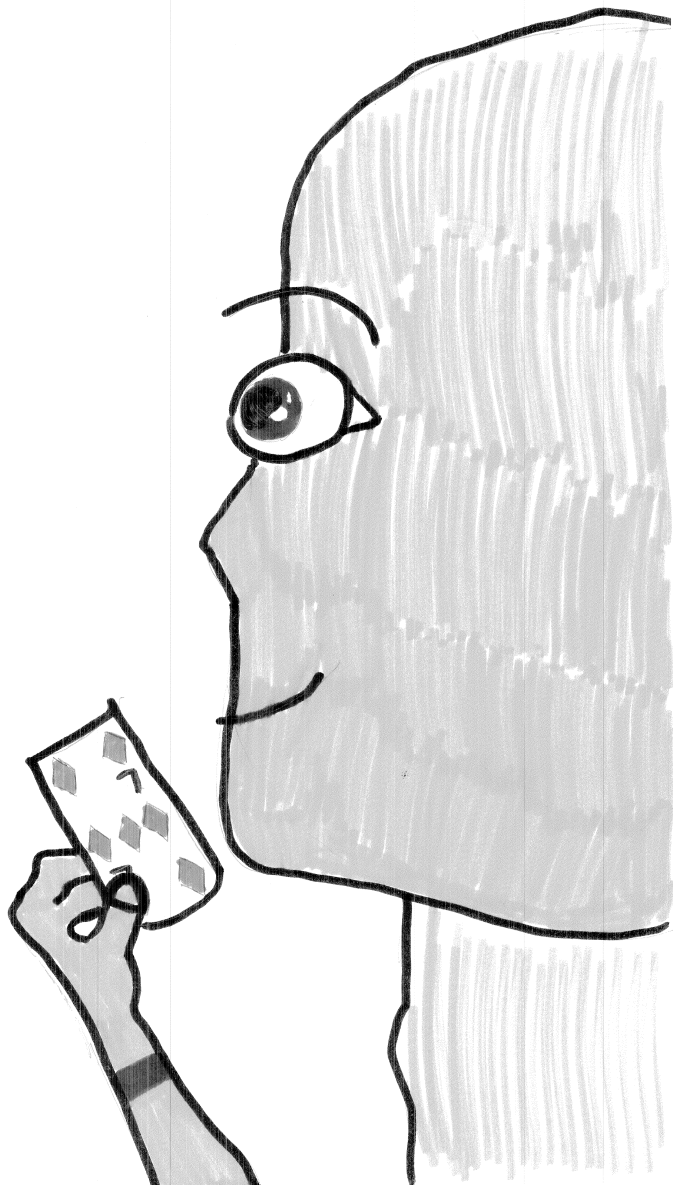
Alors, le jeune garçon se retrouva derrière de énormes barreaux, où il entendit une voix.

C'était la voix d'un autre prisonnier inconnu. Il lui a raconté que lui aussi à été enseveli par le sable, et qu'il s'est retrouvé ici.

1) Claude entendit un bruit, il a pétié.

2) Claude entendit un bruit, c'était son chien qui a fait pipi sur sa chaussette.

3) Claude entendit un bruit, mais il n'y avait pas de bruit.



La goutte d'eau

Un jour, une petite famille de fourmis marchait sur un large rocher. La veille, il avait plu. Les fourmis marchaient depuis très longtemps, en voyageant de rochers en rochers.

Mais vient l'arrivée d'une énorme goutte d'eau qui tomba du large poutier du dessus, cet incident n'était pas prévu pour les fourmis, la goutte d'eau les fit voler dans l'air pour ensuite retomber violemment dans la rivière.

Nous n'entendons plus jamais parler de ces fourmis.

GENERAL CONCLUSION

Using psychological insights seems to be appropriate in order to explain economic behaviors. Decision theory presents a great tool and answers for economics decisions-making, but it would probably benefit even more by extending the standard models by introducing psychological variables. This dissertation contributes to the literature of behavioral economics by focusing on five main psychological determinants: risk preferences, self-confidence, personality, gender and creativity. This thesis is divided in two main parts. The first part presents two chapters on the role of self-confidence on two types of decisions, in a general setting, that can be applied to educational and career choices. The second part is interested in the teenagers' behavior, specifically on schooling achievement and risk preferences. Even though each chapter has its specific goal, they all show that the chosen determinants have overall an effect on the behaviors considered.

Chapter 1 presents a laboratory experiment and a model lying on a decision to undertake a risky activity with increasing levels of difficulty. The model introduces confidence as a function of two biases: an ability-estimations bias based on the perception of ability (estimation bias) and a cognitive bias leading to over-confidence for low-ability agents, and under-confidence for high-ability agents (miscalibration bias). Hence, a psychological assumption based on two effects (the hard-easy effect and the Dunning-Kruger effect) is added to a traditional Bayesian model. Moreover, we analyze the effect of aspirations on the cognitive bias and on the decision to continue. Aspirations is considered as partly exogenous and partly endogenous. The model shows that a positive ability-estimation enhances effort if ability and effort are complementary factors of success (and the reverse if they are substitutes), but that miscalibration bias increases self-confidence but will reduce effort and success. We also show that for risk averse subjects, aspiration level has a different effect on effort according to

the ability level: a small effect on low ability agents but an increase for high ability agents. By a simple manipulation of endowments, we experimentally simulated three levels of aspiration, that we matched with the three levels of difficulty. We show that confidence and success do not interact the same way with aspirations and effort across the two ability categories of agents and we confirm the physiological assumption. A new result emerges from our analysis: we show that confidence has a good and bad component meaning that it does not always predict success. We also confirm the aspiration upgrade after a positive update of abilities. A discussion is further developed on the different possible applications (notably on education).

Concerning the decision to enter competition, we try, in chapter 2, to have a greater understanding on the fact than women enter less tournament than men. This chapter shows that confidence in one's relative performance plays a major role in the tournament entry decision. When we provide the subjects with a performance feedback, this allows subjects to update their beliefs not as Bayesian agents would, but they rather overreact to the information they receive. We also show that men and women do not process the information the same way as men seem to more internalize the information, leading them to more optimal decisions in terms of expected payoffs, whereas women overweight the feedback and are thus too much affected by it. These latter also do not enough take into account the fact that their performance can improve and that this can increase their probability to win a tournament.

Chapter 3 and 4 present a completely different approach as data was gathered in the field. They are based on a method that lies between an experiment done in the field and empirical data. We went in the classroom, we measured the creative potential of 9th grade pupils, their risk preferences with the Holt & Laury (2002) procedure, their personality with the Big five and gathered all their schooling grades. This enables us to explain schooling achievement with creativity and personality, as well as trying to explain risk aversion with these variables. There is a clear negative relationship of a

specific subtypes of creativity (verbal divergent thinking) on grades. Conscientiousness and openness also positively impact grades. Girls have higher grades in every subject but in sports. However, all things being equal, they have a lower probability to pass the middle school nation exam. This study yields questions on the role of creativity at school as well as the skills skill needed (or not) on the labor market. Furthermore, a deeper analysis better explains the gender effects we observe.

The fourth chapter shows that, even among 14 or 15 years olds, girls are already more risk averse than boys, but this could be inflected through personality. The procedure by H&L does not really yield different results from adults in terms of risk and inconsistency patterns. However, we would recommend to implement forced choices if the researcher only wants to focus risk preferences in order to limit inconsistency. One of the creativity dimensions measure (showing the ability of producing many different original and appropriate ideas) also decreases the propensity of being risk averse, which is in line with the idea that being more creative leads to taking more risks. Higher grades in scientific subjects decreases the probability of being inconsistent. Hence, cognitive abilities based on computation skills can limit this tendency. We consider the topic of this chapter important because it gives a better understanding of how teenagers react or perceive risk, and especially what shapes them. The findings can contribute to a better policy elaboration aimed for adolescents.

Confidence is endogenous and can be modulated in the lab via information, performance feedback or by implemented features like aspirations. As the first two chapters are based on experimental data, the external validity of the results may be put into questions, as well as the sample compositions that is mainly made of graduate students. Even though this thesis strongly supports the use of the experimental methods, it could be interesting to reproduce this kind of feedback procedure in the firms where tournaments are implemented or where higher hierarchical positions are considered by

the agents. It could be also useful to test the relationship between confidence, social background and abilities in the educational sphere. Especially, the French post high school educational system represents a good field of study as there are tracks of different nature: selective vs. competitive, general vs. vocational.

Moreover, the type of feedback firms or schools should send to their agents is questioned. It is perhaps better to focus on the possibility for performance to improve, rather than on a precise feedback that can have a negative value and thus discourages, refrains motivation and effort. Low and high ability agents behave differently (low ability agents tend to be overconfident but adapt better to the level of the competition, high ability tend to be underconfident but do not adapt to the level of the competition). An additional study could concern the long term effect of ability groups through peer effects, and also looking at how self-confidence varies according to the environment of mixed abilities or not. Does it enhance motivation? Does it decrease confidence bias?

The last two chapters have some shortcomings. First, the sample is rather small and is from a unique school, reducing the generalizability of the findings. It would greatly benefit from replicating this study in different types of schools to see any effect of the schools' environment and providing a greater heterogeneity for social background. Moreover, confidence elicitation could be added in order to analyze its effect on grades which could partially solve the problem of external validity of the results of the first two chapters. The extension of this work could also lead to better design education policies on the role of creativity at school, especially in France.

Personality seems to play a role in decision making and might be as well domain specific: one trait will impact one type of action, but maybe not another one. It is considered as rather stable in time and impacts risk aversion but maybe other preferences and their determinants can be studied. Some research has already shown that women are more altruistic than men or more averse to inequalities. What about the role of personality in these behaviors? Do they play any? Are the children and teenagers' re-

actions similar to adults'? This thesis only uses the Big Five questionnaire, but other personality traits could be considered such as self-control, novelty seeking (probably positively correlated with creativity), harm avoidance etc...

Although this dissertation brings new results and insights on the role played by the psychological determinants on economic behaviors and decision-making, it also generates new research questions that need to be explored.

Bibliography

- Aklin, Will M., C.W. Lejuez, Michael J. Zvolensky, and Marya Kahler, Chris W. and Gwadz.** 2005. "Evaluation of behavioral measures of risk taking propensity with inner city adolescents." *Behaviour Research and Therapy*, 43(2): 215–228. (p. 12)
- Allais, Maurice.** 1953. "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine." *Econometrica: Journal of the Econometric Society*, 503–546. (p. 11)
- Almlund, Mathilde, Angela Lee Duckworth, James J. Heckman, and Tim D. Kautz.** 2011. "Personality psychology and economics." National Bureau of Economic Research. (p. 18)
- Altonji, Joseph G.** 1995. "The Effects of High School Curriculum on Education and Labor Market Outcomes." *The Journal of Human Resources*, 30(3): pp. 409–438. (p. 128)
- Amabile, T M.** 1996. *Creativity In Context: Update To The Social Psychology Of Creativity*. Boulder, Colorado:Westview Press. (p. 129)
- Antonites, A.J., and J.J. Van Vuuren.** 2005. "Inducing entrepreneurial creativity, innovation and opportunity-finding skills." *SOUTH AFRICAN JOURNAL OF ECONOMIC AND MANAGEMENT SCIENCES*, 8(3): 255. (p. 130)
- Armantier, Olivier, and Nicolas Treich.** 2013. "Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging." *European Economic Review*. (p. 42)
- Ashton, Michael C., Kibeom Lee, Philip A. Vernon, and Kerry L. Jang.** 2000. "Fluid intelligence, crystallized intelligence, and the openness/intellect factor." *Journal of Research in Personality*, 34(2): 198–207. (p. 148)
- Askari, Muniza, Marco Gazel, and Louis Lévy-Garboua.** 2013. "Double or Quits: Self-confidence and the Intuitive Bayesian Hypothesis." (p. 45)

- Azmat, Ghazala, and Nagore Iriberry.** 2010. "The importance of relative performance feedback information: Evidence from a natural experiment using high school students." *Journal of Public Economics*, 94(7-8): 435–452. (p. 88)
- Azmat, Ghazala, and Nagore Iriberry.** 2012. "The Provision of Relative Performance Feedback Information: An Experimental Analysis of Performance and Happiness." Centre for Economic Performance, LSE CEP Discussion Papers dp1116. (p. 88)
- Baer, John, and James C. Kaufman.** 2008. "Gender differences in creativity." *The Journal of Creative Behavior*, 42(2): 75–105. (p. 144)
- Ballinger, T. Parker, and Nathaniel T. Wilcox.** 1997. "Decisions, Error and Heterogeneity." *Economic Journal*, 107(443): 1090–1105. (p. 139, 168, and 182)
- Bandura, Albert.** 1993. "Perceived Self-Efficacy in Cognitive Development and Functioning." *Educational Psychologist*, 28(2): 117–148. (p. 14, 15, and 128)
- Barber, Brad M., and Terrence Odean.** 2001. "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics*, 116(1): 261–292. (p. 20)
- Barbot, Baptiste.** 2012. *BB5 : Test de personnalité pour Adolescents*. (p. 129, 137, 143, 167, 172, and 173)
- Barratt, Ernest S.** 1985. "Impulsiveness subtraits: Arousal and information processing." *Motivation, emotion and personality*, 137–146. (p. 12)
- Batey, Mark, Adrian Furnham, and Xeniya Safiullina.** 2010. "Intelligence, general knowledge and personality as predictors of creativity." *Learning and Individual Differences*, 20(5): 532–535. (p. 22)
- Batey, Mark, Tomas Chamorro-Premuzic, and Adrian Furnham.** 2010. "Individual differences in ideational behavior: can the big five and psychometric intelligence predict creativity scores?" *Creativity Research Journal*, 22(1): 90–97. (p. 22)
- Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak.** 1964. "Measuring utility by a single-response sequential method." *Behavioral science*, 9(3): 226–232. (p. 12)

- Bénabou, Roland, and Jean Tirole.** 2002. "Self-Confidence And Personal Motivation." *The Quarterly Journal of Economics*, 117(3): 871–915. (p. 32 and 35)
- Bengtsson, Claes, Mats Persson, and Peter Willenhag.** 2005. "Gender and Overconfidence." *Economics Letters*, 86(2): 199–203. (p. 20)
- Benoît, Jean-Pierre, and Juan Dubra.** 2011. "Apparent overconfidence." *Econometrica*, 79(5): 1591–1625. (p. 16)
- Berlin, Noémi, Maud Besançon, and Jean-Louis. Tavani.** 2013. "Creativity, Personality and Schooling Achievement on Field Data." (p. 186)
- Berndt, Thomas J.** 1979. "Developmental changes in conformity to peers and parents." *Developmental Psychology*, 15(6): 608. (p. 164)
- Besançon, Maud, and Todd Lubart.** 2008. "Differences in the development of creative competencies in children schooled in diverse learning environments." *Learning and Individual Differences*, 18(4): 381 – 389. Including Special Issue on Creativity. (p. 130)
- Blavatsky, Pavlo R.** 2010. "Reverse common ratio effect." *Journal of Risk and Uncertainty*, 40(3): 219–241. (p. 139, 168, and 182)
- Blavatsky, Pavlo R., and Wolfgang R. Köhler.** 2007. *Lottery Pricing in the Becker-DeGroot-Marschak Procedure*. Inst. for Empirical Research in Economics. (p. 12)
- Blickle, Gerhard.** 1996. "Personality traits, learning strategies, and performance." *European Journal of Personality*, 10(5): 337–352. (p. 148)
- Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel.** 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources*, 43(4). (p. 128)
- Borghans, Lex, James J. Heckman, Bart H. H. Golsteyn, and Huub Meijers.** 2009. "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association*, 7(2-3): 649–658. (p. 179)

- Boudon, R.** 1973. *L'inégalité des Chances : la Mobilité Sociale dans les Sociétés Industrielles*. Paris: Armand Colin. (p. 32 and 35)
- Bourdieu, P., and J-C. Passeron.** 1964. *Les Héritiers*. Paris: Editions de Minuit. (p. 35 and 70)
- Bowles, Samuel, and Herbert Gintis.** 1975. "The Problem with Human Capital Theory-A Marxian Critique." *American Economic Review*, 65(2): 74–82. (p. 18, 26, and 128)
- Bowles, Samuel, Herbert Gintis, and Melissa Osborne.** 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature*, 39(4): 1137–1176. (p. 18, 26, and 128)
- Brand, Matthias, Elke Kalbe, Kirsten Labudda, Esther Fujiwara, Josef Kessler, and Hans J. Markowitsch.** 2005. "Decision-making impairments in patients with pathological gambling." *Psychiatry research*, 133(1): 91–99. (p. 12)
- Breen, Richard, and John H. Goldthorpe.** 1997. "Explaining Educational Differentials: Towards a Formal Rational Action Theory." *Rationality and Society*, 9(3): 275–305. (p. 36)
- Brody, Leslie R.** 1993. "On understanding gender differences in the expression of emotion." *Human feelings: Explorations in affect development and meaning*, 87–121. (p. 19)
- Brown, B. Bradford, Donna R. Clasen, and Sue A. Eicher.** 1986. "Perceptions of peer pressure, peer conformity dispositions, and self-reported behavior among adolescents." *Developmental Psychology*, 22(4): 521. (p. 164)
- Camerer, Colin, and Dan Lovallo.** 1999. "Overconfidence and excess entry: An experimental approach." *American Economic Review*, 89(1): 306–318. (p. 103 and 106)
- Cason, Timothy N., William A. Masters, and Roman M. Sheremeta.** 2010. "Entry into winner-take-all and proportional-prize contests: An experimental study." *Journal of Public Economics*, 94(9): 604 – 611. (p. 94)

- Cattell, Raymond B.** 1979. *The structure of personality in its environment*. New York : Springer Pub Co. (p. 18, 129, and 171)
- Clark, Jeremy, and Lana Friesen.** 2009. "Overconfidence in Forecasts of Own Performance: An Experimental Study*." *The Economic Journal*, 119(534): 229–251. (p. 42)
- Clifford, Margaret M.** 1991. "Risk taking: Theoretical, empirical, and educational considerations." *Educational Psychologist*, 26(3-4): 263–297. (p. 164)
- Clifford, Margaret M., , William Yun Lan, Kakanang Maneesri, Kuo-Nan Mao, and I.C. Lu.** 1991. "Gender Differences in Children's Academic Risk Taking and Failure Tolerance." (p. 168)
- Coller, Maribeth, and Melonie B. Williams.** 1999. "Eliciting Individual Discount Rates." *Experimental Economics*, 2(2): 107–127. (p. 178)
- Compte, Olivier, and Andrew Postlewaite.** 2004. "Confidence-Enhanced Performance." *American Economic Review*, 94(5): 1536–1557. (p. 14, 15, 32, and 33)
- Costa Jr, Paul, Antonio Terracciano, and Robert R. McCrae.** 2001. "Gender differences in personality traits across cultures: robust and surprising findings." *Journal of personality and social psychology*, 81(2): 322. (p. 150 and 173)
- Croson, Rachel.** 2005. "The method of experimental economics." *International Negotiation*, 10(1): 131–148. (p. 9)
- Croson, Rachel, and Uri Gneezy.** 2009. "Gender Differences in Preferences." *Journal of Economic Literature*, 47(2): 1–27. (p. 19, 86, and 168)
- Dahlbom, L., A. Jakobsson, Niklas Jakobsson, and Andreas Kotsadam.** 2011. "Gender and overconfidence: are girls really overconfident?" *Applied Economics Letters*, 18(4): 325–327. (p. 20)
- Datta Gupta, N., A. Poulsen, and M.C. Villeval.** 2012. "Gender and Competitiveness. Experimental Evidence." *Forthcoming in Economic Inquiry*. (p. 20, 86, and 116)

- Dee, Thomas S.** 2007. "Teachers and the gender gaps in student achievement." *Journal of Human Resources*, 42(3): 528–554. (p. 150)
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde.** 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review*, 100(3): 1238–60. (p. 165)
- Dollinger, Stephen J., Klaus K. Urban, and Troy A. James.** 2010. "Creativity and openness : further validation of two creative product measures." *Creativity Research Journal*, 16(1): 37–47. (p. 22)
- Drago, Francesco.** 2011. "Self-esteem and earnings." *Journal of Economic Psychology*, 32(3): 480–488. (p. 18)
- Duncan, Greg J., and Rachel Dunifon.** 2012. "Soft-Skills and Long-Run Labor Market Success." *Research in labor economics*, 35: 313–339. (p. 19)
- Duru-Bellat, M.** 2003. *Social Inequality at School and Educational Policies*. UNESCO-IIEP. (p. 35)
- Eckel, Catherine C., and Philip J. Grossman.** 2008. *Men, women and risk aversion: Experimental evidence*. Vol. 1, New York Elsevier. (p. 168)
- Eckel, Catherine C., Philip J. Grossman, Cathleen A. Johnson, Angela C.M de Oliveira, Christian Rojas, and Rick K Wilson.** 2012. "School environment and risk preferences: Experimental evidence." *Journal of Risk and Uncertainty*, 45(3): 265–292. (p. 26 and 166)
- Ellsberg, Daniel.** 1961. "Risk, ambiguity, and the Savage axioms." *The Quarterly Journal of Economics*, 643–669. (p. 11)
- Ertac, Seda.** 2011. "Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback." *Journal of Economic Behavior & Organization*, 80(3): 532–545. (p. 15, 87, 88, and 101)
- Eysenck, Sybil B.G., Paul R. Pearson, G. Easting, and John F. Allsopp.** 1985. "Age norms for impulsiveness, venturesomeness and empathy in adults." *Personality and individual differences*, 6(5): 613–619. (p. 12)

- Falch, Torberg, and Linn Renée Naper.** 2013. "Educational Evaluation Schemes and Gender Gaps in Student Achievement." *Economics of Education Review*. (p. 21, 150, and 154)
- Feist, Gregory J.** 1998. "A meta-analysis of personality in scientific and artistic creativity." *Personality and Social Psychology Review*, 2(4): 290–309. (p. 22)
- Feist, Gregory J.** 2010. "The function of personality in creativity. The Nature and nurture of the creative personality." In *The Cambridge handbook of creativity.* , ed. James C. Kaufman and Robert J Sternberg, 113–130. New York:Cambridge University Press. (p. 22)
- Feldhusen, John F, and Ban Eng Goh.** 1995. "Assessing and accessing creativity: An integrative review of theory, research, and development." *Creativity Research Journal*, 8(3): 231–247. (p. 132)
- Furnham, Adrian, Jane Zhang, and Tomas Chamorro-Premuzic.** 2006. "The relationship between psychometric and self-estimated intelligence, creativity, personality and academic achievement." *Imagination, Cognition and Personality*, 25(2): 119 – 145. (p. 130)
- Gardner, John W.** 1940. "The relation of certain personality variables to level of aspiration." *The Journal of Psychology*, 9(1): 191–206. (p. 35)
- Gardner, Margo, and Laurence Steinberg.** 2005. "Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study." *Developmental psychology*, 41(4): 625. (p. 164)
- Gervais, Simon, and Terrance Odean.** 2001. "Learning to be overconfident." *Review of financial Studies*, 14(1): 1–27. (p. 35)
- Gilboa, Itzhak, Andrew Postlewaite, and David Schmeidler.** 2009. "Rationality of belief." *Synthese*. (p. 15)
- Gneezy, Uri, Kenneth L. Leonard, and John A. List.** 2008. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *NBER Working Paper*, , (13727). (p. 20 and 86)

- Gneezy, Uri, Muriel Niederle, and Aldo Rustichini.** 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, CXVIII: 1049–1074. (p. 20 and 86)
- Goff, Maynard, and Phillip L. Ackerman.** 1992. "Personality-intelligence relations: Assessment of typical intellectual engagement." *Journal of Educational Psychology*, 84(4): 537. (p. 148)
- Goldberg, Lewis R.** 1981. "Language and individual differences: The search for universals in personality lexicons." *Review of personality and social psychology*, 2(1): 141–165. (p. 129)
- Goldberg, Lewis R., Dennis Sweeney, Peter F. Merenda, and John Edward Hughes.** 1998. "Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes." *Personality and Individual Differences*, 24(3): 393–403. (p. 18)
- Griffin, Dale, and Amos Tversky.** 1992. "The weighing of evidence and the determinants of confidence." *Cognitive psychology*, 24(3): 411–435. (p. 24 and 33)
- Grossman, Zachary, and David Owens.** 2011. "An Unlucky Feeling: Persistent Overestimation of Absolute Performance with Noisy Feedback." Department of Economics, UC Santa Barbara University of California at Santa Barbara, Economics Working Paper Series 1937484. (p. 87, 88, and 101)
- Gunther, Christina, Neslihan Arslan Ekinici, Christiane Schwierén, and Martin Strobel.** 2010. "Women can't jump?—An experiment on competitive attitudes and stereotype threat." *Journal of Economic Behavior & Organization*, 75(3): 395–401. (p. 86)
- Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin.** 2003. "Does peer ability affect student achievement?" *Journal of Applied Econometrics*, 18(5): 527–544. (p. 140)
- Harbaugh, William T., Kate Krause, and Lise Vesterlund.** 2002. "Risk Attitudes of Children and Adults: Choices Over Small and Large Probability Gains and Losses." *Experimental Economics*, 5(1): 53–84. (p. 26 and 166)

- Harrison, Glenn W., and Elisabet E. Rutström.** 2008. "Risk aversion in the laboratory." *Research in experimental economics*, 12: 41–196. (p. 12 and 178)
- Heckman, James J.** 2000. "Policies to foster human capital." *Research in economics*, 54(1): 3–56. (p. 15)
- Heckman, James J.** 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science*, 312: 1900–1902. (p. 18, 26, and 128)
- Heckman, James J., Jora Stixrud, and Sergio Urzua.** 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482. (p. 18, 26, and 128)
- Hertwig, Ralph, and Andreas Ortmann.** 2001. "Experimental practices in economics: A methodological challenge for psychologists?" *Behavioral and Brain Sciences*, 24(03): 383–403. (p. 9)
- Heunks, Felix J.** 1998. "Innovation, Creativity and Success." *Small Business Economics*, 10(3): 263–272. (p. 166)
- Hirschberger, Gilad, Victor Florian, Mario Mikulincer, Jamie L. Goldenberg, and Tom Pyszczynski.** 2002. "Gender differences in the willingness to engage in risky behavior: A terror management perspective." *Death studies*, 26(2): 117–141. (p. 19)
- Hogan, Joyce, and Brent Holland.** 2003. "Using theory to evaluate personality and job-performance relations: a socioanalytic perspective." *Journal of Applied Psychology*, 88(1): 100. (p. 18)
- Hollard, Guillaume Idriss, S?bastien Massoni, and Jean-Christophe Vergnaud.** 2010. "Subjective beliefs formation and elicitation rules: experimental evidence." Université Pantéon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne Documents de travail du Centre d'Economie de la Sorbonne. (p. 42 and 91)
- Holt, Charles, and Susan Laury.** 2002. "Risk aversion and incentive effects." *American Economic Review*, 92: 1644–1655. (p. 13, 131, 133, 138, 167, 169, and 170)

- Hoppe, Ferdinand.** 1976. "Success and failure." *Field theory as humankind: Contributions of Lewin's Berlin Group*, 454–493. (p. 34 and 47)
- Hoxby, Caroline.** 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." National Bureau of Economic Research Working Paper 7867. (p. 140)
- John, Oliver P., and Sanjay Srivastava.** 1999. "The Big Five trait taxonomy: History, measurement, and theoretical perspectives." In *Handbook of Personality: Theory and Research*. . Second ed., , ed. Lawrence A. Pervin and Oliver P. John, 102–138. New York: Guilford Press. (p. 129 and 171)
- Judge, Timothy A., and Charlice Hurst.** 2007. "Capitalizing on one's advantages: role of core self-evaluations." *Journal of Applied Psychology*, 92(5): 1212. (p. 18)
- Kahneman, Daniel, and Amos Tversky.** 1979. "Prospect Theory: An Analysis of Decisions Under Risk." *Econometrica*, 47: 263–291. (p. 11, 38, and 47)
- Kaufman, James C., and Ronald A. Beghetto.** 2009. "Beyond big and little: The four c model of creativity." *Review of General Psychology*, 13(1): 1–12. (p. 22)
- Kerr, J.H., and J. Vlaminkx.** 1997. "Gender differences in the experience of risk." *Personality and individual differences*, 22(2): 293–295. (p. 19)
- Komarraju, Meera, Steven J Karau, Ronald R Schmeck, and Alen Avdic.** 2011. "The big five personality traits, learning styles, and academic achievement." *Personality and Individual Differences*, 51(4): 472–477. (p. 129)
- Köszegi, Botond, and Matthew Rabin.** 2007. "Reference-Dependent Risk Attitudes." *American Economic Review*, 97(4): 1047–1073. (p. 47)
- Kruger, Justin, and David Dunning.** 1999. "Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments." *Journal of personality and social psychology*, 77(6): 1121. (p. 16, 24, and 33)
- Laidra, K., H. Pullmann, and J. Allik.** 2007. "Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school." *Personality and Individual Differences*, 42(3): 441–451. (p. 129)

- Lavy, Victor.** 2008. "Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment." *Journal of Public Economics*, 92(10): 2083–2105. (p. [21](#), [150](#), and [154](#))
- Lejuez, CW., Jennifer P. Read, Christopher W. Kahler, Jerry B. Richards, Susan E. Ramsey, Gregory L. Stuart, David R. Strong, and Richard A. Brown.** 2002. "Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART)." *Journal of Experimental Psychology: Applied*, 8(2): 75. (p. [12](#))
- Lent, Robert W., Steven D. Brown, and Kevin C. Larkin.** 1986. "Self-efficacy in the prediction of academic performance and perceived career options." *Journal of counseling psychology*, 33(3): 265. (p. [15](#))
- Lévy-Garboua, Louis, Hela Maafi, David Masclet, and Antoine Terracol.** 2012. "Risk aversion and framing effects." *Experimental Economics*, 15(1): 128–144. (p. [139](#), [168](#), [181](#), and [182](#))
- Lichtenstein, S., and P. Fischhoff.** 1977. "Do Those Who Know More Also Know More About What They Know?" *Organizational Behavior And Human Performance*, 20: 159–183. (p. [16](#), [24](#), and [33](#))
- Lichtenstein, S., B. Fischhoff, and LD Phillips.** 1982. "Calibration of probabilities: The state of the art to 1980." *Judgment under uncertainty: Heuristics and biases*, 306–334. (p. [16](#))
- Loewenstein, George F, Elke U Weber, Christopher K Hsee, and Ned Welch.** 2001. "Risk as feelings." *Psychological bulletin*, 127(2): 267. (p. [19](#))
- Loomes, Graham, and Robert Sugden.** 1998. "Testing Different Stochastic Specifications of Risky Choice." *Economica*, 65(260): 581–98. (p. [139](#), [168](#), and [182](#))
- Lubart, T., M; Besancon, and B. Barbot.** 2011. *EPoC : Evaluation du Potentiel Créatif*. (p. [131](#), [134](#), [135](#), [137](#), [167](#), and [173](#))
- Lubart, Todd, and Jacques-Henri Guignard.** 2004. "The generality-specificity of creativity: a multivariate approach." In *Creativity: From Potential to Realization*. ,

- ed. Robert J Sternberg, Elena L Grigorenko and Jerome L Singer, 43–56. Washington, D.C.:American Psychological Association. (p. 22)
- Lubart, Todd, Christophe Mouchiroud, Sylvie Tordjman, and Franck Zenasni.** 2003. *Psychologie de la créativité*. Paris:Armand Colin. (p. 26 and 130)
- Lubart, Todd I.** 1994. “Creativity.” In *Thinking and Problem Solving.* , ed. Robert J Sternberg, 289–332. New York:Academic Press. (p. 130)
- Maehr, Martin L., and Carol Midgley.** 1991. “Enhancing student motivation: A schoolwide approach.” *Educational Psychologist*, 26(3-4): 399–427. (p. 150)
- Majaro, Simon.** 1992. “Strategy search and creativity: The key to corporate renewal.” *European management journal*, 10(2): 230–238. (p. 130)
- McCrae, Robert R., and Paul T. Costa.** 1987. “Validation of the five-factor model of personality across instruments and observers.” *Journal of personality and social psychology*, 52(1): 81. (p. 18, 26, 129, 132, 135, 144, and 171)
- Meece, Judith L., Beverly Bower Glienke, and Samantha Burg.** 2006. “Gender and motivation.” *Journal of School Psychology*, 44(5): 351–373. (p. 150)
- Millstein, Susan G., and Bonnie L. Halpern-Felsher.** 2001. “Perceptions of risk and vulnerability.” *Adolescent risk and vulnerability: Concepts and measurement*, 15. (p. 164)
- Möbius, Markus M., and Tanya S. Rosenblat.** 2006. “Why beauty matters.” *The American Economic Review*, 222–235. (p. 15)
- Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat.** 2011. “Managing Self-Confidence: Theory and Experimental Evidence.” National Bureau of Economic Research, Inc NBER Working Papers 17014. (p. 15, 87, 88, 91, and 101)
- Montmarquette, Claude.** 2008. “L'économétrie des données expérimentales: défis et opportunités.” *Economie & prévision*, , (1): 7–17. (p. 9)

- Moore, Don A, and Paul J. Healy.** 2008. "The trouble with overconfidence." *Psychological review*, 115(2): 502. (p. 16)
- Nekby, Lena, Peter Skogman Thoursie, and Lars Vahtrik.** 2008. "Gender and self-selection into a competitive environment: Are women more overconfident than men?" *Economics Letters*, 100(3): 405–407. (p. 20)
- Niederle, Muriel, and Lise Vesterlund.** 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics*, 122(3): 1067–1101. (p. 13, 20, 86, and 89)
- Niederle, Muriel, and Lise Vesterlund.** 2011. "Gender and Competition." *Annual Review of Economics*, 3(1): 601–630. (p. 86)
- Nyhus, Ellen K., and Empar Pons.** 2005. "The effects of personality on earnings." *Journal of Economic Psychology*, 26(3): 363–384. (p. 18)
- O'Connor, Melissa C., and Sampo V. Paunonen.** 2007. "Big Five personality predictors of post-secondary academic performance." *Personality and Individual Differences*, 43(5): 971–990. (p. 129)
- Ouazad, Amine, and Lionel Page.** 2013. "Students' perceptions of teacher biases: Experimental economics in schools." *Journal of Public Economics*, 105(0): 116 – 130. (p. 21 and 150)
- Page, Lionel, Louis Levy-Garboua, and Claude Montmarquette.** 2007. "Aspiration levels and educational choices: An experimental study." *Economics of Education Review*, 26(6): 747–757. (p. 37)
- Poropat, Arthur E.** 2009. "A meta-analysis of the five-factor model of personality and academic performance." *Psychological bulletin*, 135(2): 322–38. (p. 129)
- Quaglia, Russell J., and Casey D. Cobb.** 1996. "Toward a theory of student aspirations." *Journal of Research in Rural Education*, 12(3): 127–132. (p. 34)
- Quiggin, John.** 1982. "A theory of anticipated utility." *Journal of Economic Behavior & Organization*, 3(4): 323–343. (p. 11)

- Rabin, Matthew.** 1998. "Psychology and economics." *Journal of economic literature*, 36(1): 11–46. (p. 5)
- Reyna, Valerie F, and Charles J. Brainerd.** 1995. "Fuzzy-trace theory: An interim synthesis." *Learning and Individual Differences*, 7(1): 1–75. (p. 167)
- Reyna, Valerie F., and Frank Farley.** 2006. "Risk and rationality in adolescent decision making implications for theory, practice, and public policy." *Psychological science in the public interest*, 7(1): 1–44. (p. 164)
- Richardson, Michelle, Charles Abraham, and Rod Bond.** 2012. "Psychological correlates of university students' academic performance: a systematic review and meta-analysis." *Psychological bulletin*, 138(2): 353–87. (p. 129)
- Richards, Ruth.** 1990. "Everyday creativity, eminent creativity, and health: "After-interview"? for CRJ issues on creativity and health." *Creativity Research Journal*, 3(4): 300–326. (p. 22)
- Rohrman, Bernd.** 2002. "Risk attitude scales: Concepts and questionnaires." *Melbourne: University of Melbourne*. (p. 12)
- Rubinstein, Gidi.** 2005. "The big five among male and female students of different faculties." *Personality and Individual Differences*, 38(7): 1495–1503. (p. 150 and 173)
- Runco, Mark.** 2004. "Creativity." *Annual review of psychology*, 55: 657–87. (p. 129)
- Salgado, Jesus F.** 1997. "The Five Factor Model of personality and job performance in the European Community." *Journal of Applied psychology*, 82(1): 30. (p. 18)
- Saucier, Gerard, and Lewis R. Goldberg.** 1998. "What is beyond the big five?" *Journal of Personality*, 66(4): 495–524. (p. 129, 171, and 172)
- Savage, Leonard J.** 1954. "The foundations of statistics." (p. 11)
- Sen, Arun K., and Knuta Hagtvat.** 1993. "Correlations among creativity, intelligence, personality, and academic achievement." *Perceptual and Motor Skills*, 77(2): 497–498. (p. 130)

- Shurchkov, Olga.** 2012. "Under pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints." *Journal of the European Economic Association*, 10(5): 1189–1213. (p. 86)
- Silbereisen, Rainer K, and Peter Noack.** 1988. "On the constructive role of problem behavior in adolescence." *Persons in context: Developmental processes*, 152–180. (p. 164)
- Simon, Herbert A.** 1986. "Rationality in psychology and economics." *Journal of Business*, S209–S224. (p. 5)
- Slovic, Paul, and Sarah Lichtenstein.** 1968. "Relative importance of probabilities and payoffs in risk taking." *Journal of Experimental Psychology*, 78(3p2): 1. (p. 11)
- Stankov, Lazar, Jihyun Lee, Wenshu Luo, and David J. Hogan.** 2012. "Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety?" *Learning and Individual Differences*, 22(6): 747–758. (p. 15)
- Steinberg, Laurence.** 2004. "Risk taking in adolescence: what changes, and why?" *Annals of the New York Academy of Sciences*, 1021(1): 51–58. (p. 164)
- Steinberg, Laurence.** 2007. "Risk taking in adolescence new perspectives from brain and behavioral science." *Current Directions in Psychological Science*, 16(2): 55–59. (p. 165)
- Sternberg, R.J., and T.I. Lubart.** 1995. *Defying the Crowd: cultivating creativity in a culture of conformity*. Simon & Schuster. (p. 129)
- Sternberg, Robert J., and James C. Kaufman.** 2010. *The Cambridge handbook of creativity*. Cambridge University Press. (p. 132)
- Sternberg, Robert J., and Todd I. Lubart.** 1993. "Investing in Creativity." *Psychological Inquiry*, 4(3): pp. 229–232. (p. 166)
- Sutter, Matthias, and Daniela Rützler.** 2010. "Gender differences in competition emerge early in life." (p. 20)

- Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rüetzler, and Stefan T. Trautmann.** 2013. "Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior." *American Economic Review*, 103(1): 510–31. (p. 26 and 166)
- Svenson, O.** 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica*, 47: 143–148. (p. 16)
- Tellegen, Auke.** 1982. "Brief manual for the multidimensional personality questionnaire." *Unpublished manuscript, University of Minnesota, Minneapolis*, 1031–1010. (p. 12)
- Torrance, Ellis P.** 1968. *Torrance tests of creative thinking*. Personnel Press, Incorporated. (p. 133)
- Torrance, E. P.** 1962. *Guiding Creative Talent*. Englewood Cliffs, NJ:Prentice-Hall. (p. 135)
- Trapmann, Sabrina, Benedikt Hell, Jan-Oliver W. Hirn, and Heinz Schuler.** 2007. "Meta-analysis of the relationship between the big five and academic success at university." *Journal of Psychology*, 215(2): 132–151. (p. 129)
- Tversky, Amos, and Daniel Kahneman.** 1974. "Judgment under uncertainty: Heuristics and biases." *Science*, 185(4157): 1124–1131. (p. 15)
- Tversky, Amos, and Daniel Kahneman.** 1992. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and uncertainty*, 5(4): 297–323. (p. 11)
- Van Eijck, Koen, and Paul M de Graaf.** 2004. "The Big Five at school: The impact of personality on educational attainment." *Netherlands' Journal of Social Sciences*, 40(1): 24–40. (p. 18)
- Von Neumann, John, and Oskar Morgenstern.** 1945. "Theory of games and economic behavior." *Bull. Amer. Math. Soc*, 51: 498–504. (p. 11)

- Wagner, Mervyn K.** 2001. "Behavioral characteristics related to substance abuse and risk-taking, sensation-seeking, anxiety sensitivity, and self reinforcement." *Addictive behaviors*, 26(1): 115–120. (p. 19)
- Weber, Elke U., Ann-Renée Blais, and Nancy E. Betz.** 2002. "A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors." *Journal of behavioral decision making*, 15(4): 263–290. (p. 19)
- Weiss, Andrew.** 1995. "Human capital vs. signalling explanations of wages." *The Journal of Economic Perspectives*, 9(4): 133–154. (p. 128)
- Wozniak, David, William T. Hardbaugh, and Ulrich Mayr.** 2011. "The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices." Department of Economics, UC Santa Barbara. (p. 86, 88, and 94)
- Zuckerman, Marvin, and Bernard Lubin.** 1965. *Manual for the multiple affect adjective check list*. Educational and Industrial Testing Service San Diego. (p. 134)
- Zuckerman, Marvin, and D. Michael Kuhlman.** 2000. "Personality and Risk-Taking: Common Bisocial Factors." *Journal of personality*, 68(6): 999–1029. (p. 19)
- Zuckerman, Marvin, Sybil B. Eysenck, and Hans J. Eysenck.** 1978. "Sensation seeking in England and America: cross-cultural, age, and sex comparisons." *Journal of consulting and clinical psychology*, 46(1): 139. (p. 12)

List of Tables

1.1	Initial endowments and additional gains and losses	39
1.2	Number of anagrams that have to be solved by level and curriculum . .	40
1.3	Descriptive statistics Significance levels of two-tailed z-test or t-test: ** 5% *** 1%	43
1.4	Success and failure rates, by treatment (in %)	45
1.5	Effect of the aspirations on effort depending on risk preferences.	59
1.6	Estimation of $1 - \mu$ for low-ability subjects	61
1.7	Estimation of $1 - \mu$ for high-ability subjects	61
1.8	Probit on the decision to continue to L2 and L2 success (conditional on L1 success), on low ability subjects.	65
1.9	Probit on being relegated to the low group	66
1.10	Probit on attaining middle and upper groups	67
1.11	Number of participants per session and in total	72
2.1	<i>Repetition</i> (control) and <i>Ability Group</i> (treatment) at step 4	93
2.2	Average number of correctly-solved sums at step 1 and step 2 (the p-values correspond to two-tailed t-tests).	97
2.3	LPM for the tournament entry decision of low-performing women and men.	109
2.4	LPM for the tournament entry decision of high-performing women and men.	111
2.5	LPM for the tournament entry decision among <i>Repetition</i> participants by gender.	112
2.6	LPM for the tournament entry decision among <i>Ability Group</i> participants by gender.	113
2.7	Descriptive statistics.	119
3.1	Correlations table between the four scores (p-values in parentheses). . .	136
3.2	Mandatory courses a middle school pupil attends and number of observations for each subject.	140

3.3	Correlation between the BB5 personality traits and the creativity scores (p-values in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).	143
3.4	OLS estimation on scientific subjects' grades (standardized)	147
3.5	OLS estimation on humanity subjects' grades (standardized)	148
3.6	OLS estimation on art, music and IT and sport grades (standardized)	149
3.7	Probit on the passing the final exam	152
3.8	BB5 scores by gender	160
3.9	Descriptive statistics on grades in each subject and class (A to F), on a 20 points basis	161
4.1	BB5 standard scores by gender (standard deviation in parentheses)	173
4.2	Risk aversion and proportion of safe choices	178
4.3	Interval regression on the CRRA coefficient	180
4.4	Probit on being inconsistent	184

List of Figures

1.1	Experiment scheme	41
1.2	Proportion of choices between both track according to treatments (in %).	46
1.3	49
1.4	Summary of the decision and updating process	57
1.5	Confidence levels and success rate for low ability subjects (in %). Significance levels of two-tailed z-test: * 10% ** 5% *** 1%	62
1.6	Confidence levels and success rate for high ability subjects (in %). Significance levels of two-tailed z-test: * 10% ** 5% *** 1%	63
1.7	Experiment screen faced by subjects for level 1, step 1	72
1.8	The decision process for a subject belonging to the Gain Treatment	73
2.1	Experiment summary	94
2.2	Beliefs update compared to Bayesian beliefs	100
2.3	Proportion of low-performing women (left) and men (right) choosing tournament entry in Choice 1 and Choice 2.	102
2.4	Proportion of high-performing women (left) and men (right) choosing tournament entry in Choice 1 and Choice 2.	102
2.5	Proportion of high-performing (HP) women and men entering or submitting to the tournament(in %).	107
2.6	Actual and predicted entry rates of low-performing participants by gender and treatment	115
2.7	Actual and predicted entry rates of high-performing participants by gender and treatment	116
3.1	Distribution of the four standardized creativity scores	137
3.2	Lottery choices	139
3.3	Distribution of the standardized grades in each subject	141
3.4	Creativity scores by gender	145
3.5	Two examples of GIT task.	157

3.6	Two examples of VIT task: writing a story based on the title "Drop of Water"	158
3.7	Two examples of VDT task: writing as many ends of story as possible .	159
3.8	Distribution of standardized grades per subject, by gender (O for boys, 1 for girls)	160
4.1	Lottery choices	171
4.2	Distribution of scientific and humanities scores, by gender (n=90) . . .	176
4.3	Proportion of safe choices in each decision	177
4.4	Two examples of GIT task.	188
4.5	Two examples of VIT task: writing a story based on the title "Drop of Water"	189
4.6	Two examples of VDT task: writing as many ends of story as possible .	190