



HAL
open science

Évaluer et améliorer l'efficacité des psychothérapies : adaptation française d'un système de suivi des résultats thérapeutiques et étude des processus participant au changement

Christophe Cazauvieilh

► To cite this version:

Christophe Cazauvieilh. Évaluer et améliorer l'efficacité des psychothérapies: adaptation française d'un système de suivi des résultats thérapeutiques et étude des processus participant au changement. Psychologie. Université de Bordeaux, 2018. Français. NNT: 2018BORD0261 . tel-01959442

HAL Id: tel-01959442

<https://theses.hal.science/tel-01959442v1>

Submitted on 18 Dec 2018

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.

THÈSE PRÉSENTÉE POUR OBTENIR LE GRADE DE
DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE SOCIÉTÉS, POLITIQUE, SANTÉ PUBLIQUE

Mention : PSYCHOLOGIE

Par **Christophe CAZAUVIEILH**

**Évaluer et améliorer l'efficacité des psychothérapies : adaptation
française d'un système de suivi des résultats thérapeutiques et
étude des processus participant au changement.**

Sous la direction de M. le Professeur Kamel GANA

Soutenue le 03 décembre 2018

Membres du jury :

M. le Professeur Kamel GANA
Université de Bordeaux

Directeur de thèse

M. le Professeur Jean-Louis MONESTES
Université de Grenoble

Rapporteur

M. le Professeur Bruno QUINTARD
Université de Bordeaux

Co-directeur

M. le Professeur Cyril TARQUINIO
Université de Lorraine

Rapporteur et Président du jury

Remerciements

Je n’imaginai pas en envisageant cette thèse, que je m’engageais dans une aventure qui me permettrait de rencontrer tant de gens à travers le monde, de voyager aussi bien physiquement que mentalement, et de recevoir tant d’aide à tous les niveaux. Au final, je n’imaginai pas non plus avoir tant de personnes à remercier. Je vais essayer de n’oublier personne.

M le Professeur Cyril Tarquinio et M le Professeur Jean-Louis Monestes, vous me faites un très grand honneur en acceptant de bien vouloir évaluer ce travail de thèse. Et je tiens à vous en remercier tout particulièrement.

M le Professeur Kamel Gana : je te remercie grandement d’avoir accepté d’encadrer cette thèse, de m’avoir laissé une très grande latitude d’action pour étudier et travailler sur ce que je souhaitais, de m’avoir fourni une aide statistique et des enseignements précieux, et de m’avoir guidé dans cette aventure.

M le Professeur Bruno Quintard : un très grand merci également, d’avoir accepté de co-encadrer cette thèse et d’en avoir en outre facilité toutes mes démarches. Je te remercie beaucoup pour ta confiance en me proposant notamment d’intervenir régulièrement auprès des étudiants du Master de psychologie de la santé.

Monsieur le Dr Scott D Miller : j’ai eu la chance de commencer mon travail de psychologue en bénéficiant de vos écrits et de vos enseignements. A une époque où je cherchais des outils pour m’assurer que j’apportais à mes patients des bénéfices bien réels. Votre travail a bouleversé mes quelques certitudes de l’époque. Vous avez transformé ma pratique et depuis nos premiers contacts, j’ai été passionné par vos recherches. J’ai aussi été impressionné par votre capacité à séparer le grain de l’ivraie, à synthétiser des pans entiers de la littérature sur la psychothérapie, et par votre esprit critique. Vous avez accepté que je travaille à l’adaptation de

vos outils en français et m'avez épaulé dans ce processus. Je vous adresse tous mes remerciements.

M Le Docteur Michael Lambert : vous m'avez autorisé à utiliser votre échelle dans cette recherche. Sans condition. J'ai une très grande admiration pour ce que vous avez accompli dans votre carrière et les données que vous avez collectées pendant tant d'années ont été déterminantes dans la production de ce travail. Un très grand merci.

Bill Andrews : je ne peux m'empêcher de reprendre le format de remerciements de Daryl (Chow) dans sa thèse : sans toi, ma deuxième étude auprès des psychothérapeutes aurait été tout bonnement impossible. Merci pour tout le travail accompli, pour la mise en relation, l'aide à la présentation en anglais, les crash tests. Je sais que je te dois une grande faveur.

Dr Jebb Brown : que dire ! Quel honneur de pouvoir échanger avec vous, je connaissais votre travail depuis quelques années et maintenant j'ai le plaisir d'avoir pu bénéficier de votre expertise et de vos conseils. Vous avez été d'une aide inestimable dans ce travail de thèse, notamment en ce qui concerne ma deuxième étude, mais pas que. Je vous suis extrêmement reconnaissant de tout le temps que vous avez passé à m'aider, à traquer les bugs informatiques (un très grand merci également au Dr Tak Minami dont j'ai pu utiliser le script), et à réfléchir avec moi. Je suis impatient de collaborer avec vous plus avant et je vous remercie infiniment.

M le Professeur Bruce Wampold : un très grand merci pour toute l'inspiration puisée dans votre travail et pour votre gentillesse. Je me rappelle encore la teneur de nos premiers échanges par mail, et de vos encouragements, il y a plus de dix ans, qui m'ont donné de l'élan dans mes projets de commencer à faire de la recherche. Je suis admiratif de vos compétences mathématiques et de votre savoir. Vous m'avez montré que vous étiez accessible et que les plus grands chercheurs n'étaient pas les moins accessibles. J'espère pouvoir collaborer activement avec vous dans le futur.

Un très grand merci également au Professeur Ronald Kessler, au Professeur Marc Corbières, au Professeur Franck Bond, au Professeur Robert Courtois et au Professeur Marc Covents pour m'avoir autorisé à utiliser leurs questionnaires ou leur plateforme de questionnaires pour cette étude.

Un très, très grand merci aux Dr Davison et aux Dr Yap, qui m'ont autorisé à créer une tâche de reconnaissance des micro-expressions à l'aide de leur matériel de très grande qualité.

Un très grand merci à tous mes patients, qui m'ont fait confiance et m'ont donc confié leurs données. Merci à tous les thérapeutes qui ont participé à ces études. Je leur suis extrêmement reconnaissant pour leur engagement et pour le travail accompli. Un très grand merci également, à tous les étudiants qui ont participé à cette étude.

Mme Michèle Koleck, merci beaucoup de m'avoir facilité l'accès à la recherche et à cette recherche en particulier, et de m'avoir permis de faire un premier pas dans le monde des enseignants chercheurs, en me confiant la réalisation de TD, et la dispense de cours magistraux.

Madame Chantal Durand, merci pour toutes vos informations et pour votre aide concernant les aspects juridiques de cette recherche.

Merci beaucoup à Sabrina Granger également pour toutes ses ressources, ses formations intéressantes, sa disponibilité et son aide.

Un très grand merci à mes amis, Patrice, Gerald, Laëtitia, Claire, pour leur intérêt pour ma thèse et pour leur amitié précieuse. Et pourtant, je sais que j'ai dû les embêter en évoquant constamment mon travail.

Un merci à ma famille, dont *certaines membres* ont eu l'occasion de relire *certaines* de mes productions.

Merci aussi à Monsieur T, qui ne peut pas faire entendre sa voix ici. Merci pour son attention bienveillante, sa sagesse, son équanimité et sa curiosité.

Un énorme merci à ma compagne Camille, qui a été à mes côtés à tous les instants. Son aide (multiniveau) a été une ressource considérable qui m'a permis de mener ce projet à son terme.

Et d'avance pardon à ceux que j'ai pu oublier...

Résumé

La pratique et l'encadrement des psychothérapies sont des enjeux majeurs actuels des politiques de santé publique. Même si la psychothérapie est en général une entreprise efficace et que les résultats sont comparables ou surclassent les pratiques médicales les mieux validées, les taux d'abandons, de détérioration et les mauvais résultats associés au suivi d'une psychothérapie sont loin d'être insignifiants, et les cliniciens constituent une part importante de la variance d'efficacité des soins ; sans connaître précisément cette efficacité ou sans disposer de méthodes simples permettant d'accéder à ces informations. Les cliniciens ont tendance à voir leur efficacité décliner au cours de leur carrière sans bénéficier de méthodes fiables pour assurer le développement de leur efficience.

Le Suivi en Continu des Résultats au traitement (Routine Outcome Monitoring ou ROM) est une méthodologie assez récente basée sur le recueil du feedback des patients quant à la progression et à l'alliance thérapeutique, qui sont deux prédicteurs des thérapies les plus efficaces. Cette méthodologie, bien que validée empiriquement, est peu utilisée dans les contextes de soins francophones. La pratique ROM paraît associée à de meilleurs résultats avec des patients qui ne tiraient pas initialement de bénéfices de leurs thérapies, en lien avec des gains en efficacité modérés, plutôt constatés, dans les contextes de soins externes.

Les thérapeutes ne semblent pas apprendre de leur exposition aux systèmes de feedback et les tentatives de modélisation de l'effet du feedback instrumental en psychothérapie (ROM) sont articulées autour de la modulation des effets dus au feedback par des caractéristiques plutôt liées aux thérapeutes (le genre, l'auto-efficacité, la propension à recevoir du feedback) ; les thérapeutes bénéficiant du feedback en proportions variables. Le feedback agirait ainsi comme un signal négatif correcteur qui aiderait les thérapeutes dans l'orientation du traitement délivré.

En synthétisant ces résultats, nous postulons que le feedback instrumental ROM puisse interagir avec le feedback naturel constitué de la communication non verbale présente dans toutes les séances de psychothérapie, et pour laquelle les praticiens pourraient avoir des capacités de reconnaissance variable. Le feedback instrumental viendrait ainsi compléter ou maximiser le potentiel d'utilisation de ces informations pour mieux guider la thérapie.

Dans une première étude, nous avons réalisé la validation transculturelle des échelles ORS et SRS du système PCOMS (Miller et al., 2005). Ces échelles françaises possèdent des qualités psychométriques comparables aux versions américaines en termes de structure factorielle, de validité et de fidélité. Des données prénormatives françaises, et des indices de seuils et de changement fiable sont également documentés.

Dans une seconde étude pilote, nous avons exploré l'association entre les capacités des psychothérapeutes à percevoir des signaux non verbaux (macro et micro émotions, voix et gestes) et l'efficacité due au traitement sur un échantillon de 11 psychothérapeutes anglo-saxons délivrant une psychothérapie à 386 patients en étant engagés dans une pratique ROM. Une analyse multiniveau révèle que la détection des signaux présents dans la voix (dont le contenu a été filtré par une technique de réorganisation aléatoire des segments) est associée à de meilleurs résultats. Cette donnée constitue un tout premier élément dans le sens du modèle supplétif du feedback instrumental en psychothérapie (lens model), pour lequel de nouvelles recherches sont nécessaires.

Mots-clés : Psychothérapie, Résultats, Alliance, Feedback, ROM, Communication Non Verbale

Abstract

The practice and management of psychotherapy is a current stake in health policies. Even if psychotherapy is generally deemed efficient and that outcomes compare or exceed best medical practices, drop-out, deterioration rates and poor results associated to care are far to be insignificant, and clinicians represent a major part in the variance of therapy outcomes; without having access to their results or getting simple means to access these data. Clinicians efficacy decline during the career without benefiting from reliable method to help in the growth of the efficiency.

Routine Outcome Monitoring (ROM) is a quite recent methodology based on patient feedback regarding the outcomes and the therapeutic alliance, which are predictors of the most efficient therapies. This methodology is evidence-based but not much used in French settings of care. ROM seems to be associated to better outcomes for patients who initially did not benefit from their treatments, with moderate gains, especially noticed in outpatient settings.

Therapists do not seem to learn from feedback exposure and theoretical models of ROM posit that therapist's related characteristics modulate the outcomes (gender, self-efficacy, feedback propensity); therapists vary in feedback benefits. Feedback could act as a corrective negative signal helping therapists to inform the treatment.

Summarizing these results, we suggest that ROM feedback could interact with non-verbal naturalistic feedback signals for whom therapist would vary in detection abilities. Instrumental feedback could complete or maximize the potential to use these information in order to help the therapeutic process.

In a first study we realized the French transcultural validation of the PCOMS scales (ORS and SRS, Miller et al., 2005). These French scales share many properties with the American version (factorial structure, validity, reliability). Pre-normative data, cut-off score and reliable change index are also documented.

In a second pilot study, we explored the association between therapist abilities to detect non-verbal signals (macro and micro emotion, in the voice and in the body) and treatment outcomes with a sample of 11 English speaking therapists delivering psychotherapy to 386 patients, engaged in a ROM practice. A multilevel analysis revealed that the detection of voice signals (with a randomized spliced filtering) was associated to better outcomes in treatment. This data is a first element for an auxiliary model of instrumental feedback in psychotherapy (the lens model), for which new researches are needed.

Keywords : Psychotherapy, Outcomes, Alliance, Feedback, ROM, Non verbal Communication

Unité de recherche

INSERM UMR 1219 'Bordeaux Population Health Research Center'

Equipe SEPIA "Psychoépidémiologie du vieillissement et des maladies chroniques"

3 ter, place de la Victoire

33076 Bordeaux Cedex - FRANCE

Panta rhei (Tout coule).

Heraclite

Si vous croyez savoir, vous ne savez pas.

Lao-Tseu

SOMMAIRE

1. PARTIE THEORIQUE	27
1.1. FONCTIONNEMENTS ET DYSFONCTIONNEMENTS DES PSYCHOTHERAPIES : SYNTHÈSE DE LA LITTÉRATURE ET PRÉSENTATION DE DEUX MÉTHODOLOGIES ASSOCIÉES AUX MEILLEURS RÉSULTATS	28
1.1.1. Une définition de la psychothérapie	28
1.1.2. Principaux résultats de la recherche en psychothérapie	32
1.1.2.1. Ce qui fonctionne en psychothérapie	33
1.1.2.2. Les dysfonctionnements de la psychothérapie	41
1.1.3. Comment répondre à ces enjeux ?	43
1.1.3.1. L'alliance thérapeutique	44
1.1.3.2. Les effets positifs au début du traitement	45
1.1.4. Comment intégrer et utiliser ces données pour répondre aux enjeux de la psychothérapie et guider la pratique courante	47
1.1.4.1. Le monitoring continu des résultats de la thérapie et de l'alliance (ROM)	48
1.1.4.2. L'entraînement des thérapeutes par la pratique délibérée (Deliberate Practice, Ericsson, 2006)	48
1.1.5. Les sept points clé de la psychothérapie	50
1.1.6. Synthèse des données relatives aux fonctionnements et aux dysfonctionnements des psychothérapies	53
1.2. LE SUIVI CONTINU DES RÉSULTATS EN PSYCHOTHERAPIE : REVUE DE LA LITTÉRATURE DU ROUTINE OUTCOME MONITORING (ROM) ET PRÉSENTATION DU SYSTÈME PCOMS.	54
1.2.1. Historique du Routine Outcome Monitoring	55
1.2.2. Définition et principes généraux du Suivi Continu des Résultats (ROM)	56
1.2.2.1. Les indicateurs à suivre en continu	57
1.2.2.2. Des systèmes de feedback très différents	58
1.2.3. Revue générale de la littérature sur les effets liés à la pratique ROM	61
1.2.3.1. Principaux résultats de la littérature	64
1.2.3.2. Des effets positifs pour les patients NOT (PAS EN BONNE VOIE de rétablissement)	64
1.2.3.3. Des effets positifs pour tous les patients (NOT et ON)	65
1.2.3.4. Des effets positifs quand le feedback est renvoyé au patient	66

1.2.3.5.	Une réduction des dépenses de santé _____	66
1.2.3.6.	Un effet sur la symptomatologie spécifique _____	66
1.2.3.7.	Un effet positif du au feedback instantané _____	67
1.2.3.8.	Un effet global sur le traitement _____	67
1.2.3.9.	La méta-analyse de Knaup et al. (2009) _____	67
1.2.3.10.	La revue et la méta-analyse Cochrane (Kendrick et al., 2016) _____	67
1.2.3.11.	La méta-analyse de Østergård et al. (2018) _____	70
1.2.4.	<i>Synthèse des données relatives à l'efficacité du Suivi Continu des Résultats en Psychothérapie (ROM) _____</i>	73
1.3.	VERS UN MODELE EXPLICATIF DE L'EFFET DU ROUTINE OUTCOME MONITORING : DE LA COMMUNICATION NON VERBALE AU FEEDBACK INSTRUMENTAL. _____	74
1.3.1.	<i>Données relatives à l'efficacité du Routine Outcome Monitoring dans les traitements orientés par le feedback _____</i>	74
1.3.1.1.	Les caractéristiques reliées au thérapeute dans l'utilisation du feedback _____	74
1.3.1.2.	Les thérapeutes n'apprennent pas du feedback _____	75
1.3.1.3.	Le feedback améliore parfois les résultats sans autre entraînement formel supplémentaire _____	76
1.3.2.	<i>Comment intégrer ces données et développer un modèle rendant compte de l'effet du feedback dans la ROM et dans la psychothérapie en général ? _____</i>	77
1.3.2.1.	Modèles théoriques des interventions liés au feedback _____	77
1.3.2.2.	La problématique de la rupture de l'alliance thérapeutique _____	80
1.3.2.3.	Une synthèse des modèles explicatifs du feedback dans la pratique ROM _____	81
1.3.3.	<i>De l'existence de signaux d'alerte fiables de détérioration des patients et d'une mauvaise relation thérapeutique en situation naturelle de psychothérapie _____</i>	85
1.3.3.1.	Une forme spécifique d'interaction et d'échanges verbaux et non verbaux _____	86
1.3.3.2.	La fonction des messages non verbaux _____	87
1.3.3.3.	La capacité à décoder les messages non verbaux. _____	89
1.3.3.4.	Les signaux non verbaux fiables : micro expressions et expressions subtiles _____	90
1.3.4.	<i>Vers un modèle explicatif des effets dus au feedback dans la pratique ROM _____</i>	92
1.3.4.1.	Du feedback non verbal au feedback instrumental _____	92
1.3.4.2.	Une tentative de modélisation du feedback instrumental dans la pratique ROM. _____	94
1.3.4.3.	Conjectures reliées au modèle palliatif du FI (modèle de médiation) _____	98

1.3.4.4.	Conjectures reliées au modèle supplétif du FI (modèle de modération) :	99
1.3.5.	<i>Synthèse des modèles relatifs à l'utilisation de la pratique ROM en psychothérapie :</i>	102
1.4.	SYNTHESE DE LA PARTIE THEORIQUE ET PROBLEMATIQUE DE LA RECHERCHE	103
1.4.1.	<i>Synthèse des variables d'intérêt des effets dus au feedback instrumental dans la pratique ROM.</i>	103
1.4.2.	<i>Hypothèses relatives à l'adaptation française d'un système de Routine Outcome Monitoring : le système PCOMS</i>	105
1.4.3.	<i>Hypothèses relatives aux liens entre le feedback non verbal et le feedback instrumental dans la pratique ROM.</i>	106
2.	PARTIE EMPIRIQUE	108
2.1.	ÉTUDE N°1 : VALIDATION FRANÇAISE DES ÉCHELLES DE RESULTATS ET DE L'ALLIANCE THERAPEUTIQUE ORS ET SRS (MILLER ET AL., 2005) ; L'ÉTUDE VFERA	109
2.1.1.	<i>Origines de la démarche</i>	109
2.1.2.	<i>Présentation de la méthodologie générale d'adaptation transculturelle.</i>	113
2.1.2.1.	Projet de départ	113
2.1.2.2.	Le processus d'adaptation transculturelle	114
2.1.3.	<i>Mesures utilisées</i>	119
2.1.3.1.	Le questionnaire d'impact MI-22 (Outcome Questionnaire-22, OQ-HAI-22, Lambert & Burlingame, 1996b, version française validée par Flynn et al., 2003).	119
2.1.3.2.	L'échelle Working Alliance Inventory-Short form (WAIS-S, Tracey and Kokotovic, 1989 ; version française validée par Corbière et al. (2006)).	120
2.1.3.3.	L'échelle de Kessler-6 ou K6 (Kessler-6 Scale;Kessler and al., 2003; National population Health survey, 1999 ; étude de validité française réalisée par Arnaud et al ., 2010).	120
2.1.3.4.	Le Big Five Inventory français - The Big Five Inventory (BFI, John, Donahue & Kentle, 1991 ; version française validée par Plaisant et al., 2010)	121
2.1.3.5.	Étape de pré-test des échelles.	123
2.1.4.	<i>Procédure et Participants</i>	124
2.1.4.1.	Procédure Générale	124
2.1.4.2.	Participants	125
2.1.4.3.	Critères d'inclusion et d'exclusion	128
2.1.4.4.	Échantillons issus de la population française	129

2.1.5.	<i>Analyses statistiques</i>	130
2.1.6.	<i>Validation de l'échelle Outcome Rating Scale (ORS).</i>	132
2.1.6.1.	Statistiques descriptives et données pré-normatives	132
2.1.6.2.	Analyse des items et de l'échelle	133
2.1.6.3.	Analyse de la fidélité	134
2.1.6.4.	Validité de construit	135
2.1.6.5.	Seuil Clinique et indice de changement fiable	141
2.1.7.	<i>Validation de l'Échelle d'Évaluation de la Séance-Session Rating Scale (SRS)</i>	143
2.1.7.1.	Statistiques descriptives -Données pré-normatives	143
2.1.7.2.	Analyse des items	144
2.1.7.3.	Analyse de la Fidélité	145
2.1.7.4.	Analyse de la validité du construit.	145
2.1.7.5.	Analyse de la validité critérielle	147
2.1.7.6.	Détermination d'un score seuil de l'alliance.	147
2.1.8.	<i>Discussion</i>	148
2.1.9.	<i>Conclusion</i>	152

2.2. ÉTUDE PILOTE DES RELATIONS ENTRE LA CAPACITE A DECODER LES SIGNAUX NON-VERBAUX ET LES

PERFORMANCES DES PSYCHOTHERAPEUTES, REALISEE AUPRES D'UN GROUPE DE THERAPEUTES ANGLO-SAXONS ENREGISTRANT LEURS RESULTATS DANS LE CADRE DU ROUTINE OUTCOME MONITORING : L'ÉTUDE NON VERBAL - OUTCOME (NVB-O STUDY)

154

2.2.1.	<i>Origines de la démarche</i>	154
2.2.1.1.	Synthèse de la littérature	154
2.2.1.2.	Influence des capacités relationnelles des thérapeutes	155
2.2.1.3.	Limitation des démarches de recherches antérieures	157
2.2.1.4.	Hypothèses relatives aux habiletés à percevoir les signaux non verbaux en psychothérapie.	161
2.2.2.	<i>Méthodologie générale de notre étude</i>	163
2.2.2.1.	Critères d'inclusion	163
2.2.2.2.	Procédures	164
2.2.3.	<i>Échelles administrées et données recueillies</i>	165
2.2.3.1.	Le Questionnaire Clinical Outcomes in Routine Evaluations-10	165
2.2.3.2.	Le questionnaire Patient Health Questionnaire-9	166

2.2.3.3.	Le questionnaire Generalized Anxiety Disorders -7 _____	167
2.2.3.4.	Le questionnaire Work-Related Acceptance and Action Questionnaire _____	167
2.2.3.5.	Le questionnaire Mini Profile of Nonverbal Sensitivity _____	168
2.2.3.6.	Le questionnaire Geneva Emotion Recognition Test Short Version _____	169
2.2.3.7.	La tâche de reconnaissance des Micro-expressions faciales émotionnelles (MIE-R) _____	169
2.2.3.8.	Variables contextuelles et de contrôle collectées au niveau du patient. _____	175
2.2.3.9.	Variables contextuelles collectées au niveau des thérapeutes _____	175
2.2.3.10.	Analyses statistiques _____	176
2.2.4.	Résultats _____	177
2.2.4.1.	Statistiques descriptives concernant les thérapeutes _____	178
2.2.4.2.	Statistiques descriptives concernant les patients. _____	180
2.2.4.3.	Données de résultats mesurées pendant les soins _____	181
2.2.4.4.	Données de résultats relatives aux habiletés des thérapeutes _____	182
2.2.4.5.	Analyse statistique des relations entre les capacités des thérapeutes à reconnaître les signaux non verbaux et les performances de thérapeutes. _____	186
2.2.4.6.	Analyse exploratoire des corrélations entre les scores obtenus aux questionnaires de performance dans la reconnaissance des signaux non verbaux et les scores de changement dus au traitement. _____	189
2.2.4.7.	Calcul des corrélations entre les scores moyens des prédicteurs (associations entre la reconnaissance des différents signaux non verbaux et les scores moyens de changement ajustés par la modélisation multiniveau). _____	194
2.2.4.8.	Conclusions sur les hypothèses de départ. _____	196
2.2.5.	Discussion _____	198
3.	DISCUSSION GENERALE _____	200
	REFERENCES BIBLIOGRAPHIQUES _____	209
	ANNEXES _____	231

SOMMAIRE DES FIGURES

FIGURE 1. MODELE PALLIATIF DU FEEDBACK - MODELE EN MEDIATION	95
FIGURE 2. MODELE SUPPLETIF DU FEEDBACK - MODELE EN MODERATION	96
FIGURE 3. HYPOTHESE DE L'EFFET MEDIEUR PALLIATIF DU FEEDBACK NEGATIF SUR L'ALLIANCE THERAPEUTIQUE DANS LA CONDITION NOT (MAUVAIS RESULTATS)	97
FIGURE 4. HYPOTHESE DE L'EFFET MODERATEUR / SUPPLETIF DU FEEDBACK NEGATIF SUR L'ALLIANCE THERAPEUTIQUE DANS LA CONDITION NOT (MAUVAIS RESULTATS)	98
FIGURE 5. TENTATIVE DE MODELISATION DU FEEDBACK DANS LA PRATIQUE ROM	104
FIGURE 6. SEQUENCE D'ADAPTATION TRANSCULTURELLE (INSPIRE DE VALLEYRAND, 1989)	122
FIGURE 7. ITEM A DE L'ECHELLE SRS (ETUDE VFERA)	124

SOMMAIRES DES TABLEAUX

TABLEAU 1. DONNEES NORMATIVES PRELIMINAIRES POUR L'ECHELLE ORS A LA SEANCE 1	133
TABLEAU 2. CORRELATIONS OBTENUES ENTRE LES DIFFERENTS ITEMS DE L'ORS ET LES SOUS-ECHELLES DE LA MI-22, A LA PREMIERE SEANCE DE THERAPIE EN CONDITION NON FEEDBACK. (N = 28)	136
TABLEAU 3. CORRELATIONS ENTRE LES SCORES DES ITEMS DE L'ORS ET LES SCORES DE L'ECHELLE K6 DANS LE GROUPE CLINIQUE NON FEEDBACK A LA SEANCE 1 (N = 28)	137
TABLEAU 4. CORRELATIONS ENTRE LES SCORES OBTENUS AUX ITEMS DE L'ORS, ET LES SCORES OBTENUS AUX SOUS ECHELLES DE LA MI-22 DANS LE GROUPE NON CLINIQUE AU TEST (N = 130)	138
TABLEAU 5. CORRELATIONS ENTRE LES SCORES OBTENUS AUX ITEMS DE L'ORS ET LES SCORES OBTENUS A LA K6 TOTAL DANS LE GROUPE ETUDIANT NON CLINIQUE (ENC) AU TEST (N = 130)	138
TABLEAU 6. DONNEES PRE-NORMATIVES OBTENUES A L'ECHELLE SRS DANS LE GROUPE CLINIQUE TOTAL (ECT) A LA PREMIERE SEANCE	144
TABLEAU 7. CORRELATIONS ENTRE LES SCORES A L'ECHELLE SRS (ITEMS ET SCORE TOTAL) ET LES SCORES A L'ECHELLE WAI-S (SOUS-ECHELLES ET SCORE TOTAL) EN CONDITION NON FEEDBACK A LA SEANCE 1 (N=22)	146
TABLEAU 8. SCORES BRUTS OBTENUS A L'ECHELLE CORE-10 A LA PREMIERE SEANCE (T1) ET A LA DERNIERE SEANCE DE THERAPIE DISPONIBLE (T2 ; N = 215)	181
TABLEAU 9. SCORES BRUTS OBTENUS A L'ECHELLE GAD-7 A LA PREMIERE SEANCE (T1) ET A LA DERNIERE SEANCE DE THERAPIE DISPONIBLE (T2 ; N = 171)	181
TABLEAU 10. SCORES OBTENUS A L'ECHELLE PHQ-9 A LA PREMIERE SEANCE (T1) ET A LA DERNIERE SEANCE DE THERAPIE DISPONIBLE (T2 ; N = 171)	182

TABLEAU 11. SCORES BRUTS GLOBAUX OBTENUS A LA RECONNAISSANCE DES MACRO-EMOTIONS	_____	182
(GERT-S, <i>N</i> = 11).	_____	182
TABLEAUX 12. A) ET 12. B) DES SCORES BRUTS OBTENUS A LA RECONNAISSANCE DES MACRO-EMOTIONS, PAR EMOTION (GERT-S, <i>N</i> = 11).	_____	183
TABLEAU 13. SCORES MOYENS OBTENUS A LA RECONNAISSANCE DES SIGNAUX NON VERBAUX PAR CADRANTS ET MODALITES (MINI-PONS, <i>N</i> = 11).	_____	184
TABLEAU 14. A) 14. B) SCORES MOYENS OBTENUS A LA RECONNAISSANCE DES SIGNAUX NON VERBAUX PAR MODALITES (MINI-PONS, <i>N</i> = 11).	_____	185
TABLEAU 15. SCORES BRUTS OBTENUS A LA RECONNAISSANCE DES MICRO EXPRESSIONS (MIE-R, <i>N</i> = 11)	__	185
TABLEAU 16. CORRELATIONS DE PEARSON (R) ENTRE LES SCORES DE CHANGEMENT STANDARDISES-AJUSTES OBTENUS AUPRES DES PATIENTS ET LES SCORES DE RECONNAISSANCE DES SIGNAUX NON VERBAUX OBTENUS PAR LES THERAPEUTES. ÉCHANTILLON INITIAL TOTAL (<i>N</i> = 386)	_____	191
TABLEAU 17. : CORRELATIONS POUR LES SCORES CORE-10 (<i>N</i> = 215)	_____	192
TABLEAU 18. : CORRELATION POUR LES SCORES PHQGAD (<i>N</i> = 171)	_____	194

Introduction

Les difficultés psychologiques sont désormais reconnues comme des facteurs associés à de très importants effets négatifs sur la santé : la pratique et les recherches sur les psychothérapies représentent, en conséquence, des enjeux majeurs de santé publique mondiale. Ainsi, la dépression est classée par l'Organisation Mondiale de la Santé au premier rang des causes d'incapacité, tout en rappelant que les traitements psychologiques font partie des réponses qui peuvent être apportées en première intention. La dépression apparaîtrait en ce sens comme une épidémie « galopante » avec plus de 300 millions de personnes affectées dans le monde et une augmentation entre 2005 et 2015 de plus de 18% des personnes impactées¹.

Bien que des réserves soient émises quant au caractère endémique de cette progression, réserves qui peuvent être en particulier retrouvées dans les propos du Professeur Allen Frances, ancien architecte du DSM-IV², considérant que ce phénomène illustre une volonté des entreprises pharmaceutiques de « création ou de façonnage de la maladie » (en anglais « disease mongering »), la souffrance psychologique ne semble pas diminuer à notre époque et apparaît même protéiforme.

En parallèle, alors que notre civilisation connaît un essor technologique considérable, l'appauvrissement et la destruction des liens sociaux vont paradoxalement croissants. C'est pourtant le tissu social et la capacité de coopération qui a permis aux êtres humains au fil des siècles de survivre et de croître en conquérant les territoires plus ou moins hostiles de notre planète. Ce tissu social semble dorénavant se rétrécir comme une peau de chagrin, du moins

¹ <http://www.who.int/fr/news-room/detail/30-03-2017--depression-let-s-talk-says-who-as-depression-tops-list-of-causes-of-ill-health>

² Allen Frances est apparu récemment dans le documentaire de Michèle Dominicci « Dépression une épidémie mondiale » : https://boutique.arte.tv/detail/depression_une_epidemie_mondiale

dans les pays occidentaux. Les réseaux sociaux, qui ont connu eux aussi un essor fulgurant, apparaissent comme les « ersatz » informatiques de nos relations et ne peuvent les remplacer. Il est maintenant bien documenté que cet isolement qui correspond à la restriction objective du réseau social, et la partie subjective de ce phénomène (le sentiment associé de solitude), sont eux aussi reliés à de graves problèmes de santé et à un risque de mortalité accru (Umberson, & Montez, 2010). Dans une revue méta-analytique investiguant les liens entre la solitude, l'isolement social et la mortalité, Holt-Lunstad et al. (2015) ont ainsi mis en évidence que le manque de connections est effectivement associé à un risque plus élevé de mortalité prématurée. Ce risque est même apparu comparable à des facteurs déjà bien identifiés et responsables du développement de problématiques graves de santé, telles que les maladies cardio-vasculaires, à savoir : le surpoids, le manque d'activité physique ou les addictions. De manière plus étonnante, les auteurs de cette étude n'ont pas pu mettre en évidence de différence d'impact sur la santé en considérant le sentiment de solitude, ou l'isolement social à proprement parler ; ceci pourrait suggérer que la perception du manque de « nourritures affectives » puisse avoir un effet potentiellement délétère sur la santé, comparable aux situations de précarité relationnelle. Cet isolement qui diffère selon les âges, est plus particulièrement marqué chez les personnes âgées (Gest-emerson & Jayawardhana, 2015), et pour en prendre conscience, il suffit par exemple de se remémorer le bilan de l'épisode de canicule qui s'est abattu sur la France en 2003. Ce dernier a laissé entrevoir dans quelle précarité relationnelle pouvaient se retrouver certains de nos aînés ; et les conséquences gravissimes qui pouvaient y être associées.

La psychothérapie fait pourtant clairement partie des éléments de réponses à apporter à ces enjeux de santé publique. Il existe en effet plusieurs formes de traitement de la dépression validées de longue date (Elkin et al.1989), et répertoriées dans les listes de traitements basés sur les preuves ; les effets de la psychothérapie peuvent être comparés à ceux des traitements médicamenteux (Cuijpers, et al. 2013), avec l'avantage de meilleurs résultats à long terme

(Sugarman, 2016). En fonction du degré de sévérité de l'épisode dépressif, les agences de santé européennes recommandent que la psychothérapie constitue la première ligne de traitement³.

La psychothérapie peut également aider les individus à mieux créer, entretenir, et vivre des relations interpersonnelles satisfaisantes ; en facilitant le bon développement des habiletés relationnelles (comme la capacité de communiquer, d'exprimer ses besoins, ses ressentis, de mettre des limites). Et ce n'est donc pas un hasard si la thérapie interpersonnelle dont l'objet est de permettre au patient de développer des relations proches satisfaisantes, est reconnue comme un traitement efficace de la dépression (Elkin et al., 1989). Les êtres humains sont des êtres de sens, mais aussi avant tout, des êtres relationnels qui se construisent et se développent à partir des relations prototypiques parentales et familiales. Si celles-ci sont suffisamment sécurisées, elles permettent de se protéger, et de croître, c'est-à-dire de s'individualiser et de s'orienter dans le monde selon ses besoins et aspirations en faisant face aux difficultés de la vie. Si ces « proto-relations » présentent des carences, elles nécessitent des expériences correctrices⁴, qui peuvent être véhiculées par les expériences de vie et/ou par l'intervention d'un tiers de confiance, pour choisir (sélectionner) et apprendre (renforcer) de nouvelles façons plus fonctionnelles d'être au monde et d'être à soi. Et mieux prendre soin de soi, peut également amener à éviter ou à mieux traiter des maladies dites « somatiques » et à réduire le fardeau des problématiques de santé, au niveau d'une population.

En ce sens l'arsenal psychothérapeutique disponible, même si la métaphore peut paraître légèrement guerrière, renferme des solutions concrètes à même d'impacter la santé des individus, mais aussi, à même d'impacter le système de santé d'une population prise dans son ensemble, en réduisant les coûts liés aux problématiques de santé que la psychothérapie peut

³Pour le royaume uni voir : <https://www.nice.org.uk/guidance/cg90>, pour la France voir : https://www.has-sante.fr/portail/upload/docs/application/pdf/2017-10/depression_adulte_fiche_de_synthese_pec.pdf

⁴ Pour reprendre l'expression d'Alexander et French, 1946 – « Expérience émotionnelle correctrice »

soit traiter (prévention tertiaire), soit prévenir (prévention primaire et secondaire). Dans les pays anglo-saxons et particulièrement aux États-Unis, l'impact positif de la psychothérapie sur les coûts de santé, supportés par la société, a été documenté par Nicholas Cummings, dans les années 1950, et a donné lieu à la prise en charge des traitements par les systèmes d'assurance-santé. En Europe, les décideurs politiques sont de plus en plus sensibles à ces arguments, tout comme les organismes remboursant les soins de santé.

En France depuis quelques années, des initiatives au départ isolées de remboursement des séances de psychothérapie, comme par exemple le partenariat de la Mutuelle Carte Blanche avec les praticiens de l'Association de Thérapie Comportementale et Cognitive (AFTCC) ou le remboursement des séances de psychothérapie par la MGEN, se multiplient. Et deux initiatives actuelles nous semblent particulièrement importantes car elles sont susceptibles de stimuler la réflexion et la prise de décision quant à la prise en charge du coût des psychothérapies : une première initiative de la députée Michèle Delaunay visant à permettre l'accès aux soins délivrés par les psychologues exerçant en libéral, aux enfants, aux adolescents, et aux jeunes adultes (de 6 à 21 ans), en couvrant les frais de séances de psychothérapie (Conrath & Houdry, 2017). Une deuxième initiative plus récente, de test à grande échelle sur trois départements de la prise en charge des difficultés psychologiques par la psychothérapie pour les adultes. Ces initiatives ont initialement suscité autant d'enthousiasme que d'interrogations, parce que les dispositifs en question s'inscrivent d'office dans un parcours médicalisé, et parce que les psychologues portant le titre de psychothérapeute qui est maintenant protégé par un décret, revendiquent une indépendance d'exercice relative à leur domaine d'expertise. Le corps médical pris dans son ensemble, doute en effet, à tort selon nous, des capacités des psychologues cliniciens à détecter les difficultés psychologiques, à réaliser une bonne évaluation et à adopter une thérapie efficace, voire à référer les patients à un médecin spécialiste en cas de nécessité.

Ces initiatives vont impliquer de disposer de stratégies pour faciliter l'accès aux soins mais aussi et d'outils fiables pour encadrer leur délivrance en s'assurant que les patients retirent

des bénéfices réels des séances de psychothérapie. Et vont également amener à des recherches complémentaires, sur le terrain, pour évaluer leur coût/efficacité. A terme, l'espoir d'un accès facilité aux bénéfices de la psychothérapie pour les patients tout venant, peut maintenant être embrassé.

Pourtant, après plus d'un siècle d'exercice de la psychothérapie dite « moderne », il serait malvenu d'affirmer que nous savons comment et quand la psychothérapie fonctionne. Et même que nous savons à quelles occasions elle dysfonctionne significativement. Le rapport de l'INSERM paru en 2004 sur l'évaluation de l'efficacité des trois principales approches psychothérapeutiques représentées en France illustre bien ce point. Ce rapport fut tout aussi bien critiqué sur la forme que sur le fond. Avec des interrogations quant à la méthode de l'enquête qui nous semblent tout à fait pertinentes, car elles autorisent la remise en question des interprétations basées sur des données qui peuvent faire défaut, ou permettent au contraire d'en assurer les acquis. Et une attaque sur le fond, qui nous paraît difficilement défendable : le psychisme et les effets dus aux psychothérapies ne pourraient pas, de par leur nature, être étudiés à la lueur des méthodes expérimentales, et en conséquence l'évaluation serait impossible. La publication du rapport de l'INSERM a été suivie d'affrontements entre les différents courants, écoles de psychothérapie ou « chapelles », entre les différents chercheurs spécialistes de la question, qui pour nous, ont sonné comme le rappel à l'ordre quant à l'aspect limité des connaissances relatives aux psychothérapies. Ce signal a participé en son temps à une réflexion sur notre propre pratique naissante de psychothérapeute, et plus tard à l'initiation de ce travail de thèse.

De très nombreuses questions continuent donc à être posées, sans qu'il soit possible de dégager un consensus de toutes les réponses : que savons-nous de la pratique de la psychothérapie, de son individualisation face aux pratiques « traditionnelles » de soin ? La psychothérapie est-elle réellement efficace dans les études de laboratoire, mais aussi dans les séances dispensées par les praticiens de terrain ? Dans quelle mesure ? Comment se situe-elle

par rapport à l'efficacité des prises en charge médicales ? Est-elle efficace pour un ensemble de difficultés psychologiques ou plutôt pour des problématiques bien circonscrites ? Pose-t-elle, comme tout traitement, des problèmes d'effets inattendus, et dans quelle proportion ? Est-elle généralement, facile à suivre, ou provoque-t-elle des abandons conséquents ?

Bien d'autres questions d'intérêt peuvent également être posées : certaines psychothérapies sont-elles *toujours* plus efficaces que les autres, et ce pour quels troubles ? Comment se situent les psychothérapeutes, sont-ils généralement *tous* efficaces, est-ce qu'ils ont connaissance de leur niveau d'efficacité ? Leur efficacité est-elle liée à des formations ou des pratiques spécifiques ? La psychothérapie est-elle plus efficace de nos jours qu'il y a cinquante ans ? Comment l'évaluer ? Et comment au final l'améliorer ?

Il ne nous sera pas possible, bien entendu, de traiter extensivement de toutes ces questions dans ce travail de recherche, sachant que chaque thématique pourrait aisément faire l'objet d'un travail de thèse sans en épuiser son contenu. Aussi, nous nous situerons dans un travail de synthèse de l'information disponible en visant un objectif avant tout pragmatique, à savoir, la provision d'éléments de réponse à quatre grandes questions, que nous explorerons dans deux grandes parties.

Une première partie théorique visant à répondre à deux questions :

1) Quelles sont les connaissances actuellement disponibles quant à ce qui fonctionne et dysfonctionne dans la pratique de la psychothérapie ? Considérer cette question nous amènera notamment à résumer les principaux résultats de la recherche sur l'efficacité des psychothérapies, et à identifier les indicateurs liés aux meilleurs résultats et reliés au développement de l'efficacité des cliniciens.

2) Comment évaluer et améliorer l'efficacité des psychothérapies ? Cette question trans-théorique et trans-diagnostique, nous conduira à synthétiser, dans un premier temps, les données de la recherche sur les systèmes de gestion des résultats de la psychothérapie, utilisant le recueil

du feedback des patients en pratique courante ; et permettant à la fois d'évaluer et d'améliorer la réponse des patients à leur traitement en continu (Routine Outcome Monitoring ou ROM). Dans un deuxième temps, nous essaierons de cerner quels sont les facteurs associés à l'efficacité des interventions utilisant ces systèmes de feedback, afin d'aider au développement de thérapies plus efficaces. Nous concluons ensuite cette première partie en nous appuyant sur des données transdisciplinaires, qui nous permettront de proposer un modèle d'analyse original de l'effet des interventions basées sur le feedback en psychothérapie. En effet, il est documenté que l'utilisation de ces systèmes de suivi continu de la thérapie puisse améliorer l'efficacité des traitements, mais les mécanismes participant au changement ne sont pas clairement explicités ou modélisés.

La deuxième partie empirique de ce travail, sera articulée autour de deux dernières questions :

3) Comment adapter ces connaissances pour les rendre disponibles et utilisables par la communauté de praticiens francophones ? Cette question d'abord pragmatique nous amènera à la première adaptation et validation en français, d'un système complet de suivi continu des résultats de la psychothérapie utilisable en pratique courante, qui est reconnu comme empiriquement validé par les autorités de santé aux Etats-Unis (SAHMSA). Les systèmes de monitoring continu des psychothérapies, bien que disponibles et assez fréquemment utilisés dans les pays anglo-saxons, sont à notre connaissance moins habituellement utilisés dans les contextes francophones. En outre, il n'existe pas, à notre connaissance, de système complet de suivi des résultats de la psychothérapie, adapté pour la pratique clinique courante, qui ait été validé avec la population française.

4) Enfin, nous nous interrogerons sur comment mieux comprendre et expliquer l'efficacité de ces systèmes de feedback et par extension comment mieux expliquer une partie de l'efficacité de la psychothérapie. En nous appuyant sur le modèle théorique générique que nous aurons précédemment développé, nous essaierons de mieux comprendre une partie des

processus au cœur des changements produits par les traitements orientés par le feedback, en analysant à l'aide d'un dispositif original, utilisant à la fois des questionnaires d'évaluation éprouvés, et une nouvelle tâche d'analyse des performances des thérapeutes créée pour cette étude, les données de résultats de séances de psychothérapie obtenues par une cohorte de psychothérapeutes dans leur pratique clinique.

En orientant notre travail autour de l'étude des systèmes de feedback et de la validation française d'un système de Routine Outcome Monitoring, nous espérons amener des éléments concrets de réponse à ces enjeux de santé publique : d'abord aux cliniciens, qui sont ainsi en première ligne pour aider leurs patients ; ensuite aux « décideurs » qui ont un besoin pressant de méthodes et d'informations fiables, de données probantes relatives à l'encadrement et à la délivrance des psychothérapies dans notre pays et dans les régions francophones ; et enfin peut-être aux chercheurs, en stimulant l'initiation de nouvelles recherches sur l'importance du feedback en psychothérapie.

1. PARTIE THEORIQUE

1.1. Fonctionnements et dysfonctionnements des psychothérapies : synthèse de la littérature et présentation de deux méthodologies associées aux meilleurs résultats

Il est tout d'abord nécessaire de situer le phénomène de la psychothérapie dans une perspective chronologique et scientifique avant de l'analyser, d'en retenir une définition qui puisse être l'objet d'un consensus, puis de parcourir les revues les plus récentes sur ce qui fonctionne et dysfonctionne en psychothérapie. Nous allons ainsi dans un premier temps tenter de dégager les grandes problématiques liées à l'efficacité des psychothérapies.

Afin de produire une synthèse de la littérature disponible et d'éclairer des territoires inconnus, nous procéderons par questionnement en nous interrogeant sur sept points principaux liés à l'évaluation et à l'efficacité des psychothérapies.

1.1.1. Une définition de la psychothérapie

Il convient dans un premier temps de circonscrire le phénomène des psychothérapies et de le situer dans une perspective diachronique. En effet, la psychothérapie est une discipline tirant ses racines de pratiques ancestrales de guérison (comme par exemple le chamanisme ; Frank & Frank, 1993 ; Wampold, 2001a) ; même si elle est clairement l'objet d'investigation par l'utilisation de la méthode expérimentale depuis le 20^{ème} siècle. Les racines historiques dans l'investigation de l'efficacité des psychothérapies en France nous renvoient au 18^{ème} siècle, où Franz Anton Mesmer devint célèbre en inventant une forme de thérapie préfigurant l'hypnotisme (le "Mesmérisme"). Cette pratique était censée véhiculer un fluide spécifique ("le magnétisme animal") et soigner de nombreux problèmes médicaux. Sur ordre du roi Louis XVI, un comité scientifique dirigé par Benjamin Franklin réalisa une des premières études expérimentales d'une intervention psychothérapeutique, qui conduisit à la première démonstration de l'implication d'un processus psychologique dans les effets dus au « traitement » : le rôle de la suggestion et de l'imagination. (Cautin in Norcross, 2011).

Le caractère de scientificité n'est apparu que relativement récemment en regard de l'histoire de l'humanité. Toutes les civilisations ont en effet eu recours à des praticiens dont la

vocation était aussi bien de soigner le corps que « l'esprit », et à notre époque l'ethnopsychologie et l'ethnopsychiatrie éclairent comment ces multiples pratiques continuent à s'insérer dans un contexte culturel. L'idée qu'un dialogue stratégiquement mené puisse être curatif est particulièrement ancienne et peut notamment être retrouvée chez Socrate (469-399 AJC ; Carr, 2007) préfigurant les thérapies dites « modernes » comme les Thérapies Cognitivo-Comportementales (TCC).

Le champ des psychothérapies est de ce fait très vaste, et la recension de Lambert (2013) nous donne une idée du nombre d'écoles, ou de courants différents, de méthodes qui peuvent être potentiellement dénombrées (plus de 400 à la date de parution de l'ouvrage en 2013). Une myriade d'interventions aux contenus pseudo-scientifiques gravitent également autour de la pratique de la psychothérapie.

Aussi, nous avons décidé d'explicitier la définition du psychologue Américain Bruce Wampold, qui, selon nous, a le mérite de reprendre les éléments qui de nos jours sont communément acceptés comme concourant à la délimitation du phénomène que nous appelons psychothérapie, et permet d'en opérationnaliser ses principaux concepts :

« La Psychothérapie est principalement un traitement interpersonnel qui est basé sur des principes psychologiques et qui implique un thérapeute formé et un client qui a un trouble mental, un problème ou une plainte ; le thérapeute s'attend à ce qu'il soit (traitement) correcteur pour le trouble du client, son problème, ou sa plainte ; et celui-ci est adapté ou individualisé pour ce client particulier et son trouble, problème, ou plainte » (Wampold, 2001b, *The great psychotherapy debate*, p. 3).

Une première dimension se dégage de cette définition et définit la nature du phénomène (son essence). La psychothérapie est avant tout une entreprise relationnelle, ce qui revient à affirmer que les thérapies psychologiques sont basées sur une unité minimale constituée par la rencontre de deux individus (la dyade comme unité d'interaction). Les formes de psychothérapie basées sur des interactions et des présentations à priori plus « instrumentales »,

et dont le caractère relationnel n'est plus fondé sur la présence de deux individus ne sont pas totalement exclues de cette définition ; nous pouvons citer à ce sujet la biblio-thérapie ayant démontré son efficacité ; ou également l'utilisation de plus en plus fréquente de thérapies délivrées par un ordinateur ou par des applications (par exemple sur smartphone) ; tout en notant que dans ces formes de thérapie, la thérapie procède à une forme de relation avec le patient, ou opère un retour vers cette relation ; par exemple dans les thérapies par avatar virtuel, en développement, pour laquelle une tentative de modélisation de la relation patient-thérapeute par une intelligence artificielle, est réalisée. La thérapie de groupe, les thérapies de couple et les thérapies familiales ne sont évidemment pas exclues de cette définition, puisque la relation entre chaque participant et le thérapeute renvoie à cette unité minimale de sens, à laquelle viennent s'additionner les interactions entre chaque participant.

La deuxième dimension importante de cette définition concerne les moyens opératoires, la thérapie reposant en effet sur l'existence de principes psychologiques, de stratégies et de techniques qui fondent les bases de l'intervention et définissent différents niveaux d'abstraction (respectivement des aspects les plus théoriques aux aspects les plus techniques). Dans le travail de Wampold, cette notion participe à l'opérationnalisation de ce qu'il appelle thérapie « *Bona fide* » ou « thérapie de bonne foi », que nous pourrions traduire par « thérapie rationnelle », non pas au sens cartésien du terme, ne s'agissant pas ici de qualifier la psychothérapie d'entreprise « raisonnable », mais du fait qu'elle repose sur un *rationnel* de traitement, un critère indispensable permettant de distinguer les psychothérapies des interventions qui ne reposent pas forcément sur un corpus de connaissances accepté par la communauté scientifique (par exemple les pratiques de certains psycho-praticiens qui peuvent s'appuyer sur un corpus dont les fondements scientifiques n'ont pas été questionnés).

La troisième dimension se rapporte à la formation (ou l'entraînement du thérapeute) qui doit nécessairement être un professionnel du soin pour garantir des compétences de base et une bonne éthique.

La quatrième dimension fait référence à l'objet de la psychothérapie, la présence d'un trouble mental ou d'un problème spécifique et reprend le principe des classifications nosographiques, tout en élargissant le champ définitoire de la psychothérapie aux problèmes et plaintes expérimentées par les patients au cours de leur vie. Cette dimension existentielle, correspondant bien à la réalité du terrain où, sans exclure les patients présentant des troubles parfaitement distincts et pouvant faire l'objet d'études contrôlées et randomisées (par exemple une phobie spécifique les qualifiant de « patients purs » pour cette problématique), l'accent est mis sur le fait que les praticiens « traitent » dans la réalité des patients présentant de multiples problématiques (Kupfer et al., 2002).

La cinquième dimension, sociocognitive par nature, qualifie les croyances du côté du thérapeute, et fait référence à l'attente de résultats du praticien quant aux effets du traitement (et aux croyances socialement acceptées par la communauté scientifique). Ces effets, nous le verrons, représentent une part significative dans la variance des résultats dus au traitement.

La sixième dimension est à proprement parler contextuelle, et répond à l'impératif d'individualisation du traitement au patient particulier reçu par le praticien. Particulièrement importante dans le travail de Wampold, cette dimension permettrait de distinguer la psychothérapie de la médecine en opposant une pratique basée sur un modèle contextuel, et répondant à la question « Est-ce que cette relation de soin/thérapie est efficace, dans ce contexte de vie particulier, avec ce patient particulier » à la pratique reliée au modèle biomédical classique qui questionne plus volontiers sur « Quel est le traitement le plus adapté en fonction des symptômes présentés par ce patient ». Et nous retrouvons ici les recommandations d'individualisation du traitement éditées par l'Association Américaine de Psychologie sur les meilleures pratiques en psychothérapie (APA Presidential Task Force on Evidence-Based Practice, 2006).

1.1.2. Principaux résultats de la recherche en psychothérapie

Le point de départ du questionnement moderne sur l'efficacité des psychothérapies peut être ramené selon nous à l'article princeps d'Eysenck (1952) passant en revue 24 études suggérant que la psychothérapie n'était pas efficace et que le passage du temps surclassait l'efficacité d'une prise en charge (dans la rémission des névroses). Bien que souffrant de nombreux biais méthodologiques (groupes non comparables, critères d'améliorations questionnables), cette étude a eu pour fonction de stimuler le champ de la recherche en psychothérapie, qui a atteint depuis des proportions considérables. Lambert (2013) évoque ainsi le chiffre de quelques 60000 recherches académiques ayant donné lieu à publication, sur les 30 dernières années, tout en soulignant l'entreprise démesurée que représenterait la lecture d'une telle somme (environ 13 ans à raison de 12 heures par jour), et suggérant que devant l'impossibilité de traiter toute cette information, il est maintenant nécessaire de dégager des lignes directrices qui font consensus et d'encourager le débat sur les résultats qui ont été jusque-là acquis comme des évidences. Par ailleurs, avec l'hyper-connexion de notre société et les nouvelles possibilités offertes par l'intelligence artificielle quant à l'exploitation des données (Big Data), ce mouvement n'est pas près de s'essouffler. Aussi, il est rapidement apparu nécessaire d'utiliser une métrique commune pour rassembler les différentes informations extraites de ces études.

Cette métrique appelée taille de l'effet (en anglais, effect-size) correspond à la magnitude du changement évaluée par une échelle entre deux occasions de mesure (voir Cohen, 1988). Cette métrique étant une mesure standardisée, c'est-à-dire sans unité, elle autorise des comparaisons entre des études différentes, mais aussi des comparaisons entre des champs de recherche disciplinaires différents (par exemple la psychothérapie et les prises en charge médicales) et offre l'avantage de pouvoir agréger les données pour synthétiser les résultats. La taille de l'effet permet tout aussi bien de mesurer le changement avant-après sur une cohorte de patients recevant un traitement psychothérapeutique que d'évaluer la différence d'efficacité

entre deux traitements différents reçus par deux groupes de patients différents ; ou entre un traitement et un groupe contrôle, avec une meilleure capacité discriminative si les patients sont répartis de manière aléatoire entre les groupes, du fait de la diminution de la variance non systématique. En tant qu'indice de changement, la taille de l'effet peut être sommée et permet par exemple d'agréger les données et de déterminer les effets d'un traitement dans un corpus de recherches dédié à ce traitement ou à une problématique particulière (Meta-analyse et Mega-analyse) ; elle permet également de procéder au test des différentes composantes d'un traitement en les introduisant progressivement ou en les enlevant de manière séquentielle (procédure de démantèlement ou « Dismantling »). La puissance de l'étude, définie par la capacité à mettre en évidence la présence d'un effet s'il existe réellement, diffère selon les séquences expérimentales (« experimental design »), avec un ordre croissant de puissance des revues de la littérature, des études contrôlées (randomisées ou non) aux Méta et Méga-analyses et aux études de Démantèlement. Les Méta-analyses et les Méga-analyses ont ainsi permis de mettre en évidence des tendances bien définies sur ce qui fonctionne et ne fonctionne pas en psychothérapie en utilisant les données issues de plus de 50 ans de recherche, et d'en réaliser l'inventaire.

1.1.2.1. Ce qui fonctionne en psychothérapie

L'efficacité (efficacy) des psychothérapies

La première question qui fut posée par l'étude d'Eysenck concernait l'efficacité réelle des psychothérapies. Il fait maintenant consensus que la psychothérapie est une technique d'intervention en général véritablement efficace, en comparaison avec une absence de traitement ou avec l'administration d'un placebo (Smith, Glass, & Miller, 1980) ; une taille de l'effet moyenne de .80 (effet fort selon Cohen, 1988) est rapportée dans de nombreuses études quantitatives, quand il s'agit de comparer le fait de recevoir un traitement psychologique à l'absence de traitement pour une problématique donnée (Lambert & Ogles, 2004). Pour illustrer ce résultat, nous pouvons effectuer une comparaison imaginaire entre un individu qui aurait

reçu un traitement et un groupe d'individus qui n'auraient pas reçu de traitement pour une même problématique identifiée. L'individu traité se trouverait amélioré par rapport à près de 79% des individus qui n'auraient pas reçu de traitement face aux mêmes difficultés. Ce résultat indique que l'efficacité générale de la psychothérapie (efficacy) est forte. Cet effet est d'autant plus important si nous le comparons, ce que permet la mesure de la taille de l'effet, aux effets attribuables aux interventions médicales les plus sophistiquées : les résultats sont égaux ou supérieurs à la taille d'effet de nombreux traitements médicaux reconnus (effet moyen des interventions médicales = .05 ; Caspi, 2004 ; Cooper, 2008). Il est à noter que classiquement l'efficacité d'une psychothérapie est évaluée dans des conditions expérimentales rigoureuses. Pour permettre de calculer cette efficacité, et de tenter d'attribuer les effets aux seules vertus du traitement, il est effectivement nécessaire de contrôler au mieux toutes les variables parasites, ou biais, pouvant impacter les résultats finaux. Cet impératif relié à la démarche scientifique expérimentale, permet de rendre peu ou prou les conditions de délivrance entre les individus en « toutes choses égales par ailleurs », à l'exception de la variation systématique du traitement ou d'une autre forme d'intervention. Les critères d'inclusion et d'exclusion des patients sont donc particulièrement stricts (basés sur des critères nosologiques précis), les contingences structurelles de délivrance du traitement doivent être identiques pour tous les patients, avec la même séquence de traitement suivant un manuel unique pour chaque type de traitement comparé, le même nombre d'heures d'interventions dispensées, la même formation, supervision et expérience du clinicien, et la répartition aléatoire des patients à l'intérieur des différents groupes (intervention ou contrôle) qui est censée lisser les différences interindividuelles susceptibles d'influencer les résultats au-delà du type de traitement prodigué. La possibilité de masquer le type d'intervention délivrée au patient, pour coller au plus près des protocoles d'évaluation des traitements médicamenteux (où les praticiens et les patients sont en *double* aveugle), est censée également contrôler les effets d'attente ou placebo du côté des patients. Il est évidemment impossible pour le praticien d'être « aveuglé » quant à l'intervention

qu'il délivre, et les effets d'attentes ou de croyance (allégeance) dans le traitement ne peuvent donc pas être contrôlés de manière totalement satisfaisante par ces dispositifs expérimentaux.

L'efficience (effectiveness) des psychothérapies

La deuxième question à se poser concerne de ce fait la potentielle différence dans les résultats dus au traitement dans les conditions artificielles du laboratoire (efficacy ou efficacité) et dans les conditions naturelles de la pratique clinique courante (effectiveness ou efficience). Nous l'avons vu, dans les conditions rigoureusement contrôlées des études de laboratoire les patients sont réputés « purs », c'est-à-dire ne cotant par exemple que pour une problématique répertoriée au DSM, les thérapeutes sont spécialistes du trouble à traiter, extrêmement entraînés et supervisés à dispenser le traitement – et suivent des conditions très rigoureuses de standardisation et d'administration comme une thérapie suivant fixement un manuel de traitement. La pratique de terrain est toute autre, puisque les patients peuvent présenter de multiples problématiques et que les thérapeutes varient en termes d'expérience, de formation, de supervision, et dans leur adhérence à délivrer le protocole de traitement. Un consensus émerge pourtant quant au fait que les résultats de terrain puissent être comparables en termes de changement cliniquement significatif, en conditions réelles (voire par exemple la Méta-analyse de Shadish et al., 2000 ; Stiles et al., 2008). Il est également admis que la psychothérapie est efficace à tous les âges de la vie (Campbell et al. 2013).

La différence d'efficacité entre les psychothérapies

Une troisième question concerne la différence d'efficacité entre les différents écoles ou courants de la psychothérapie. Cette question, hautement polémique, a été le terrain d'affrontement entre les partisans de la spécificité des psychothérapies, qui argumentent que des « ingrédients » spécifiques ou composantes de traitement spécifiques propres à chaque psychothérapie sont responsables de la variance de changement due au traitement, et les partisans de l'approche dite des facteurs communs qui soutiennent que des ingrédients ou facteurs communs présents en proportions variables dans toutes les approches

psychothérapeutiques seraient responsables de la variance due au traitement. Cette question a été mise en avant en France, lors de la publication du rapport de l'Inserm sur l'efficacité de trois approches évaluées (approche psychodynamique, TCC, thérapie familiale et systémique. Inserm, 2004) dont les principaux résultats aboutirent à un classement par ordre d'efficacité basé sur une sommation du nombre d'études mesurant l'efficacité des trois orientations testées, à partir de la littérature disponible. Avec un tableau de résultat classant par ordre croissant d'efficacité : les thérapies psychodynamiques, la thérapie familiale et systémique et enfin les TCC⁵. Ce rapport fut contesté aussi bien pour des raisons légitimes incluant des problèmes méthodologiques que pour des présupposés éthiques qui sont eux difficilement défendables. Un des principaux écueils à éviter est en effet le classement d'efficacité basé sur le comptage des études soutenant l'efficacité d'une approche en particulier, amenant en fonction du nombre de résultats comptés à un grade de preuve d'efficacité supérieur. Pour illustrer cette méthodologie, nous pourrions classer les résultats de deux sportifs qui ont courus un nombre différent de marathons sur des terrains différents, dans des conditions climatiques différentes, et donc avec des variables parasites importantes non contrôlées, sur leur nombre respectif de victoires cumulées, au lieu de les faire concourir l'un avec l'autre directement au cours de plusieurs épreuves. Cette méthodologie de classement est évidemment parfaitement discutable et ne rend pas aussi bien compte de la réalité du terrain qu'une comparaison d'approches deux à deux pour une problématique donnée, qui dans le meilleur des cas peut faire l'objet d'une méta-analyse et donner l'occasion de synthétiser avec suffisamment de puissance statistique un pan de la recherche en psychothérapie.

⁵ Pour une synthèse de ces résultats voir :

http://www.ipubli.inserm.fr/bitstream/handle/10608/147/expcol_2004_psychotherapie_synthese_fr.pdf?sequence=1

Les prétendus présupposés éthiques associés au discrédit du rapport de l'Inserm, ont été liés aussi bien à l'aspect de trivialité que représenterait la « mesure » du psychisme ou de la subjectivité qu'à son impossibilité opératoire et technique ; des raisons peu soutenables tenant peut-être plus d'un tabou français. Ce discrédit ne servant pas au final ceux qui devraient l'être, c'est-à-dire les patients qui reçoivent de soins et peuvent légitimement avoir le choix et s'attendre à de vrais résultats mesurés.

La réplication des résultats des études antérieures par la méta-analyse de Smith et Glass (Smith, Glass & Miller, 1980), laissait clairement apparaître que les thérapies avaient en général des effets équivalents quand elles étaient comparées entre elles, pour des problématiques variées.

Une autre manière de répondre à cette question est donc de procéder à des méta-analyses incluant uniquement des comparaisons de recherches deux à deux de traitements *Bona-fide*. Cette démarche aboutit à un résultat qui fait à la fois consensus mais qui est souvent mal interprété et entraîne de fortes résistances de la part des partisans de la spécificité dans l'efficacité des psychothérapies : le verdict du dodo.

Ce verdict est une métaphore basée sur la phrase de la course de Dodo initialement empruntée par le psychologue Paul Rozenweig (Bertolino, Bargmann & Miller, 2012) à l'œuvre de Lewis Carroll, *Alice aux pays des merveilles*. Dans cet ouvrage, à l'issue d'une compétition, l'oiseau Dodo proclame que « Tout le monde a gagné et que tout le monde doit avoir un prix » ; le verdict du Dodo désigne à ce jour un des résultats les plus répliqués de la littérature, mais aussi un des résultats à plus haut potentiel de polémique que nous souhaitons intituler « équivalence générale des effets dus aux psychothérapies ».

Contrairement à l'interprétation classique qui peut en être faite, l'équivalence générale d'efficacité des psychothérapies ne revient pas à affirmer que pour n'importe quel patient donné et pour n'importe quelle problématique les traitements sont interchangeables puisque d'efficacité identique (« tout se vaudrait »), aboutissant à un relativisme radical, mais qu'au

contraire, dans une population donnée de patients consultants pour un trouble donné, il y aura une répartition des effets dus au traitement telle qu'il ne sera pas possible d'affirmer à priori avec certitude qu'une intervention *bona-fide* spécifique, c'est-à-dire qui est censée être thérapeutique (efficace) pour ce trouble donné, sera toujours supérieure aux autres interventions *bona-fide*. En effet, l'équivalence générale d'efficacité des psychothérapies est un résultat statistique obtenu auprès d'un patient « moyen ». Cela revient donc à dire qu'il faut individualiser, selon le contexte, pour chaque patient le traitement *bonafide* en s'assurant que celui qui est dispensé produit bien une réponse positive ; d'où l'intérêt du Suivi Continu des Résultats en cours de psychothérapie (SCR, ou Routine Outcome Monitoring - ROM) permettant de vérifier la réponse au traitement et d'adapter le traitement aux caractéristiques des patients. Ces données impliquent deux éléments extrêmement importants : cette comparaison exclut les « pseudo-traitements » qui n'ont pas démontré la preuve de leur efficacité pour une problématique donnée (« tout ne se vaut pas ») ; et il est nécessaire de suivre (monitorer) la réponse du patient au traitement en temps réel, afin d'individualiser sa délivrance (Groupe de Travail de l'APA : APA Task Force on Evidence-Based Practice, 2006, pp. 273, 276-277). De manière plus spécifique, Wampold (2001b) estime que la différence d'efficacité entre les traitements *Bonafide*, représenterait une taille d'effet maximale de .2, ce qui est considéré comme un effet faible. Les coûts d'implémentation du « meilleur traitement » ne justifieraient pas le déploiement (systématique) en comparaison de l'efficacité supplémentaire obtenue ; et en contrôlant les variables parasites telles que l'allegiance des thérapeutes, c'est-à-dire la croyance dans l'efficacité des effets que la thérapie peut produire ; cette différence de taille de l'effet spécifique serait nulle (Wampold & Imel, 2015). Les données disponibles suggèrent qu'il n'a pas été trouvé de différences dans les résultats attribuables au traitement pour la psychothérapie en général (Wampold et al., 1997) ; mais également qu'il n'existerait pas de différence dans la prise en charge de nombreux troubles spécifiques (Bertolino, Bargmann & Miller, 2012) : par exemple dans le traitement de la dépression (Wampold et al.,

2002), dans la prise en charge du stress post traumatique (Benish, Imel, & Wampold, 2008), pour les troubles liés à la consommation d'alcool (Imel et al., 2008) et pour les quatre catégories diagnostiques les plus représentées chez les enfants et les jeunes (la dépression, le trouble d'attention-hyperactivité, l'anxiété, les troubles des conduites ; Miller, Wampold, & Varhely, 2008). Une étude plus récente (Werbart et al., 2013) a exploré dans le système de soins psychiatriques publics en Suède, la différence d'efficacité due à différents traitements (TCC, approche psychodynamique, approche éclectique intégrative) : bien que la majorité des patients aient tirés des bénéfices significatifs des soins, il ne fut pas trouvé de différence significative dans l'efficacité due aux différents traitements sur trois mesures différentes de résultats. Enfin, sans prétendre couvrir tout le champ des problématiques pouvant être étudiées à la lueur de l'effet Dodo, nous pouvons citer une dernière étude méta-analytique avec comparaison directe de traitements, où il n'apparaissait pas, une fois de plus, de différence marquée dans l'efficacité due aux différents traitements (efficacité différentielle) pour des patients traités pour des troubles alimentaires (Grenon et al., 2018).

Le verdict du Dodo amène donc à questionner la part des « ingrédients spécifiques du changement (« la technique psychothérapeutique ») et la part des ingrédients présentés par toutes les psychothérapies *Bona-fide* (classiquement dénommés « facteurs communs »).

La désormais classique estimation des effets dus au traitement, issue de la recension de la littérature de Lambert, laissait apparaître que les facteurs non spécifiques représentaient environ 85% de la variance due au traitement (Asay & Lambert, 1999). Il est nécessaire de considérer que ce partitionnement de la variance des résultats en fonction d'ingrédients quantifiables et discrets, permis par l'analyse statistique, ne remet pas en cause le caractère interactionnel des processus à l'œuvre dans le traitement. Ainsi, dans les 13 à 20 % de la variance de changement des patients estimée être due au traitement, il est permis de préciser le poids relatif des différentes composantes (Bertolino, Bargmann & Miller, 2012, p6) : la technique/le modèle de psychothérapie représenterait seulement 1% de la variance due au

traitement alors que la part due au thérapeute serait, elle, de 4 à 9%. L'alliance thérapeutique représenterait à elle seule de 5 à 8 % de la variance des résultats, et les effets d'attente chez le thérapeute (allégeance) et chez le patient (placebo) rendraient compte d'environ 4 % de la variance due au traitement. Il apparaîtrait que le poids de la relation thérapeutique serait plus important que celui du type de traitement donné et que la variance dans l'efficacité des psychothérapies serait plus liée aux thérapeutes qui dispensent la technique et le traitement qu'à la technique en elle-même, sans oblitérer la nécessité d'une structure dans la thérapie et l'utilisation d'une thérapie *bona-fide*. Ce résultat très important oriente le débat sur l'évaluation et l'amélioration de l'efficacité des psychothérapies, en précisant sur qui et sur quoi doivent être recentrés les efforts d'investigation et d'amélioration.

Les effets de la psychothérapie

Une quatrième question peut être posée en relation avec la nature des effets de la psychothérapie. Il est en effet largement démontré que la psychothérapie réduit un éventail très varié de symptômes et de problématique de santé et permet de revenir à un niveau de fonctionnement normal dans de nombreuses problématiques (Lambert, 2013). Il est également clair qu'au-delà du rétablissement, suivre une psychothérapie procure des gains complémentaires en permettant de développer des compétences, et des stratégies de coping qui pourront être utilisées dans d'autres domaines de vie (Lambert & Ogles, 2004). En outre, le rapport efficacité sur coût de la thérapie est favorable, la psychothérapie diminuant les coûts de santé. La psychothérapie réduit la durée des séjours des patients, l'utilisation de soins de médecine générale et les coûts des médicaments associés (Altman et al, 2018 ; Chiles, et al., 1999 ; Kraft, et al., 2006). Une étude récente a permis de transposer ces résultats à la France, estimant que pour un euro investi dans le traitement psychothérapeutique, le ratio coût-bénéfice varierait selon la symptomatologie entre 1,14 euros et 1,95 euros (Dezetter, et al., 2014).

Aussi, cette première partie de l'analyse de la littérature brosse un portrait globalement très positif de la psychothérapie, il est de ce fait indispensable de considérer des données plus

négatives qui rendent une illustration plus réaliste, bien qu'en demi-teinte, de son efficacité réelle.

1.1.2.2. Les dysfonctionnements de la psychothérapie

La cinquième question qui peut être posée interroge ainsi sur les effets potentiellement négatifs de la psychothérapie, leurs proportions, et leurs conséquences.

La problématique des effets négatifs

Comme dans tout traitement, il existe des effets non-attendus et des effets secondaires qui peuvent être considérés comme relativement faibles, mais non négligeables. Pour en donner une estimation, une enquête récente de Crawford et al. (2016) documente des effets négatifs persistants pour 1 personne sondée sur 20 (5 %). Une classification des pratiques potentiellement dangereuses a été entreprise (et inclut notamment des techniques comme le Rebirth) mais elle reste cependant assez limitée (Lilienfeld, 2007). De manière plus préoccupante, une proportion non négligeable de patients voit leur état se détériorer pendant le traitement : 5 à 10 % (Cooper, 2008 ; Lambert & Ogles, 2004). Les données sont encore plus inquiétantes chez les enfants pour qui les taux de détérioration pourraient atteindre de 12 à 20 %, suggérant que pour une partie de cette population, les traitements ne sont clairement pas efficaces (Warren et al., 2010).

La problématique des abandons de traitement

L'abandon des soins qui peut être opérationnalisé par de multiples indicateurs (ne pas continuer la thérapie jusqu'à épuisement total des effets bénéfiques, abandon ou arrêt non planifié des séances) est malheureusement un phénomène couramment constaté par les psychothérapeutes et le taux d'abandon des soins est réputé élevé. Il diffère selon la méthodologie de la recherche employée (aux alentours de 25% pour Swift & Greenberg, 2012 ; jusqu'à 47% dans la méta-analyse de Wierzbicki & Pekarik, 1993). Pourtant, il est clair que la réponse psychothérapeutique suit une loi dose-effet, et que l'abandon précoce des séances,

même chez les patients qui ont initialement retiré des bénéfices du traitement, ne leur permet pas de retirer tous les bénéfices qui auraient pu être expérimentés au cours de la thérapie (Howard et al., 1986).

Les cliniciens n'évaluent pas leurs résultats de manière fiable

La sixième question dont nous allons traiter concerne l'habileté des cliniciens à évaluer leur efficacité durant la délivrance de la psychothérapie. Les psychothérapeutes aussi formés, expérimentés et entraînés soient-ils, ne sont pas à l'abri de présenter des biais cognitifs qui peuvent affecter les décisions prises en cours de traitement. La part de variance due aux thérapeutes est importante, et pourtant les cliniciens ont une tendance marquée à surestimer leurs capacités et leur habileté à détecter les patients qui sont à risque de se détériorer ou d'abandonner le traitement prématurément (Lambert & Ogles, 1997).

L'étude princeps de Hannan et al. (2005) est particulièrement éclairante sur ce sujet. Il s'agissait de déterminer les capacités des psychothérapeutes (22 professionnels et 22 en formation) à détecter dans leur patientèle, les personnes à risque de se détériorer pendant le traitement et de comparer leur sensibilité à la détection à celle d'un algorithme prédictif de l'échec au traitement. S'inscrivant dans le champ des recherches centrées sur le patient, et utilisant un questionnaire de suivi des résultats en continu (SCR) aux qualités psychométriques démontrées (L'OQ-45), un pool initial de patients de cette population a été sondé, autorisant à calculer un pourcentage moyen de détérioration attendu sur cette population. Armé de ces données, un algorithme de prédiction de l'échec au traitement permit de déterminer, en fonction du score actuel (de l'évolution actuelle) d'un patient sur le questionnaire OQ-45, si ce patient ressemblait plus aux patients qui se sont par le passé détériorés, ou si au contraire les données actuelles laissaient envisager une issue toute positive au traitement. La question posée aux thérapeutes qui ne bénéficiaient pas du retour d'information (feedback instrumental) de l'algorithme était de déterminer quels patients allaient réellement se détériorer.

Afin de les aider, le taux moyen de détérioration précédemment calculé et attendu sur cette population leur était communiqué. Le résultat est sans équivoque. Sur les 550 patients de l'étude, l'algorithme a identifié correctement 36 des patients qui se sont réellement détériorés pendant le traitement et les cliniciens ont identifié un seul patient sur les 40. Ceci laisse apparaître que les thérapeutes ont de grandes difficultés à identifier dans une population de patients ceux qui sont à risques de mauvais résultats, et un premier élément de réponse à ces enjeux peut être apporté par ces méthodes actuarielles d'algorithme prédictif, rendu possible par le Suivi en Continu des Résultats du traitement (SCR - ROM).

Enfin, *une septième question* résume à elle seule l'ambiguïté ou le paradoxe des recherches sur l'évaluation et l'amélioration de la psychothérapie. Quel est l'évolution de l'efficacité des psychothérapeutes au fil de leur carrière ? Là aussi, les éléments de réponse sont plutôt décourageants. Sans efforts spécifiques, en dépit des formations continues de plus en plus sophistiquées, de la supervision, et indépendamment de l'accumulation de l'expérience clinique, les thérapeutes atteignent un plateau dans leur efficacité assez rapidement au cours du temps (Bertolino, Bargmann & Miller, 2012), certaines études suggèrent même que les performances déclinent pendant la carrière (Goldberg et al., 2016).

1.1.3. Comment répondre à ces enjeux ?

Des éléments de réponse se dégagent pourtant nettement dans la littérature, mais les considérer revient à s'écarter de l'orientation qui a été donnée au champ de la psychothérapie, au cours de ces vingt dernières années. Cette orientation était définie par une stratégie « translaturée » du champ biomédical et des modèles basés sur les preuves, et s'articulant autour de l'étude et de la dissémination de thérapies « empiriquement soutenues » ou empiriquement validées (EST) ; avec un focus constant sur la fidélité dans la délivrance de la psychothérapie (manuel de traitement) et sur la formation classique et la supervision des thérapeutes, censée maximiser la bonne délivrance des ingrédients spécifiques aux traitements. Cette stratégie n'a pourtant pas démontré une réelle efficacité (Wampold & Imel, 2015).

Les éléments de réponse sont pourtant, à la fois trans-théoriques, trans-diagnostiques et même trans-disciplinaires, et concernent deux indicateurs systématiquement associés à de meilleurs résultats thérapeutiques, qui sont retrouvés dans les publications et qui permettraient de donner une nouvelle orientation pour la formation et pour la délivrance des soins (Cazauvieilh, 2013, 2014, 2015, annexe 1). En effet, deux prédicteurs sont associés à l'efficacité en psychothérapie, *une bonne alliance thérapeutique* entre le patient et le thérapeute *et des résultats positifs* sur les premières séances de traitement.

1.1.3.1. L'alliance thérapeutique

Concept bien connu des praticiens et initialement issu du courant psychodynamique, la notion d'alliance thérapeutique revêt des significations diverses. Nous pourrions utiliser le synonyme de *rapport collaboratif* (entre le patient et son thérapeute) ; mais nous pouvons également distinguer une définition qui fait l'objet d'un consensus. Dans cette acception, l'alliance thérapeutique serait composée de trois dimensions en interactions (Bordin, 1979), qui sont respectivement la dimension relationnelle (une entente ou la qualité du lien entre le patient et le thérapeute), une dimension reliée aux objectifs de la thérapie (l'entente sur les buts) et une dimension reliée à la méthode ou technique employée (entente sur les tâches).

Plusieurs caractéristiques importantes se dégagent de cette découverte, répliquée dans plusieurs milliers d'articles de recherche : la capacité du thérapeute à former des relations avec les patients prédit les résultats (Baldwin, Wampold & Imel, 2007 ; Horvath & Bedi, 2002). Une des critiques opposées à ce constat consiste à souligner la nature corrélationnelle de la démonstration. Pourtant, l'alliance thérapeutique prédit bien l'engagement dans le traitement et l'engagement est lui-même associé aux meilleurs résultats ; l'alliance étant de ce fait, un des meilleurs prédicteurs des résultats (Orlinsky, Rønnestad & Willutzki, 2003). Il existe des disparités dans les évaluations de l'alliance réalisées chez les thérapeutes et chez leurs patients, et ce sont les évaluations faites par le patient quant à l'alliance qui sont les plus prédictives du succès au traitement (Bachelor et Horvath, 1999). L'effet estimé dû à l'alliance dans les

résultats diffère selon les études, mais il est estimé que 54% de la variance des résultats dus au traitement pourrait être expliquée par l'alliance thérapeutique (Wampold, 2001). Dans une des plus importantes études contrôlées randomisées sur le traitement de la dépression (le programme de recherche collaboratives du NIMH ; Elkin et al., 1989) comparant les effets d'une thérapie comportementale, aux effets dus à la thérapie interpersonnelle et à la prise d'antidépresseur, face au placebo, l'alliance apparut comme un prédicteur du succès dans toutes les conditions expérimentales. Il est à noter que le verdict du dodo a été retrouvé dans cette étude dont le design expérimental était pour le moins extrêmement rigoureux.

Une rapide comparaison des effets dus à l'alliance dans la part de variance due au traitement, face à la taille d'effet due à la technique utilisée, laisse apparaître que l'alliance expliquerait sept fois plus de variance des résultats (Bertolino, Bargmann & Miller, 2012). L'alliance est en outre un concept transdisciplinaire et des données récentes soulignent son importance en médecine et en psychopharmacologie (Baudrant-Boga, Lehmann & Allenet, 2012 ; Consoli, 2006 ; Ireps, 2012 ; Pringle & al., 2011 ; Pringle & Colley, 2015). Ces données soulignent l'intérêt de suivre la qualité de la relation thérapeutique tout au long du traitement pour mieux engager le patient dans sa thérapie.

1.1.3.2. Les effets positifs au début du traitement

Il s'agit d'une variable moins connue et beaucoup moins considérée, bien que potentiellement opérationnalisable, dans l'amélioration de la qualité des soins. Des résultats positifs au début du traitement, en termes de changement perçu subjectivement, prédisent la réussite à l'issue des soins et la probabilité de l'issue positive au traitement diminue avec chaque séance ne présentant pas de manifestation de changement (Lambert & Ogles, 2004).

Il est nécessaire de rappeler ici les résultats de l'étude classique sur la relation dose-réponse ou l'effet-dose en psychothérapie, mis en évidence par la méga-analyse de Howard & al. (1986), à partir d'un pool de 15 bases de données de recherches couvrant une population de

plus de 2400 patients suivis en psychothérapie, et pour lesquels il s'agissait de modéliser statistiquement la réponse au traitement en fonction du nombre de séances attendues. À la 8ème séance, quasiment 50% des patients étaient améliorés de manière quantifiable et approximativement 75% à la 26ème séance (6 mois de traitement). Il existerait une diminution des gains tirés de la thérapie au cours du temps, telle que sur les premières séances, si nous matérialisons les gains par un graphique d'évolution symptomatique (par exemple en mesurant un niveau de détresse psychologique), nous pourrions constater la présence d'une forte pente descendante, indiquant un haut niveau de diminution de la symptomatologie, et tel que cette « accélération du changement » aurait tendance à diminuer au fil de séances. Ceci implique que certains patients s'améliorent très rapidement. Depuis ces résultats préliminaires, des études complémentaires ont rapporté un même schéma général d'un changement rapide suivi par des gains plus faibles au fur et à mesure de l'avancement du traitement, tout en soulignant que les patients les plus sévèrement en détresse nécessitaient plus de séances. Par exemple, l'analyse de survie de Anderson et Lambert (2001) rapporte une médiane de 11 séances pour atteindre un changement cliniquement significatif, et une nécessité de 8 séances supplémentaires chez les patients plus sévèrement perturbés. Certains auteurs suggèrent tout de même que l'efficacité de la thérapie n'est pas dose-dépendante (Stiles et al., 2008). Enfin, tous les patients, ne suivent pas une réponse linéaire au traitement et des sous-groupes de « répondeurs » ont pu être identifiés. De même, le changement peut être modélisé en phase plutôt qu'en simple relation entre la dose et l'effet. Nous pouvons citer à ce sujet l'étude de Stulz, et al. (2007), utilisant une modélisation du développement latent (Growth Mixture Model) sur les 6 premières séances de patients recevant des consultations externes, et identifiant cinq patterns ou groupes de changements correspondant aux patients et basés sur leur niveau initial de détresse (mesuré par le questionnaire CORE) ; ces groupes présentaient des réponses différentes au traitement, tout en présentant des changements en phases aux caractéristiques communes : une amélioration du bien-être rapide, suivie d'une diminution des symptômes psychologiques puis d'une

amélioration plus lente du fonctionnement. Pour éclairer ce point, nous pouvons citer les données issues des recherches sur les algorithmes de prédiction de la réussite au traitement (utilisant par exemple le questionnaire OQ-45) : le signal d'alerte quant à un mauvais résultat attendu au traitement peut apparaître avec ce système de gestion des résultats, aussitôt qu'à la troisième séance avec aux alentours de 40% de la variance des résultats prédits à l'issue du traitement (Lambert & Hansen, 2001 ; Whipple et al., 2003). Ainsi, en utilisant d'importantes bases de données normatives, Brown et Lambert (1998) découvrirent que le degré de sévérité initiale combiné au changement précoce (changement après une session de psychothérapie) prédisait l'issue du traitement (avec 17% de la variance des résultats). En découle une philosophie du traitement qui se centre sur la distribution plus importante d'attention aux patients qui ne tirent pas de bénéfices réels de leur traitement. Ces données, même si elles présentent des disparités, soulignent que le changement peut et doit être monitoré, dès le début du traitement et en continu pour s'assurer de la réponse du patient et ajuster le traitement.

1.1.4. Comment intégrer et utiliser ces données pour répondre aux enjeux de la psychothérapie et guider la pratique courante

Comme le soulignent Wampold & Imel (2015), dans une synthèse des résultats de la recherche, les arguments soutenant le modèle biomédical classique des effets de la psychothérapie sont minces, et les arguments en faveur d'un modèle contextuel de la psychothérapie sont plus nombreux. Les auteurs recommandent donc de se centrer sur ce qui est responsable du bon « fonctionnement » des psychothérapies plutôt que de continuer à investir des fonds en quantité considérable dans les essais contrôlés randomisés visant à nourrir les listes de traitements empiriquement validés. Cet agenda les amène à considérer naturellement un paradigme de recherche relativement récent, à même de réduire l'écart entre la pratique et la recherche en psychothérapie, le courant de la recherche basée sur la pratique (Practice Based Evidence ; voir Duncan et al., 2010) qui utilise les données issues de la pratique (notamment les progrès des clients), pour améliorer la qualité des soins. En considérant que ces

données sont aussi probantes que les données recueillies dans les protocoles de recherche en laboratoire, ce protocole aboutit à un renversement de la preuve, sans contester les acquis de la recherche.

1.1.4.1. Le monitoring continu des résultats de la thérapie et de l'alliance (ROM)

L'utilisation du feedback des patients en thérapie pour informer le praticien (parfois le patient également) et guider le processus thérapeutique est un exemple direct de ces recherches (que nous pouvons regrouper sous l'appellation Traitement Orienté par le Feedback ou TOF ; pour une revue voir Cazauvieilh, 2013, 2014, 2015). Ce Suivi Continu des Résultats (ROM) est un développement prometteur des recherches centrées sur le patient (Lambert & Hansen, 2001). Une recension récente de la littérature de Schuckard et al. (2016) documente que la pratique ROM peut amener dans certaines études à une diminution des taux d'abandon (jusqu'à 50%), réduire les taux de détérioration pendant les soins (jusqu'à 33%), et peut être efficace en maximisant le changement cliniquement significatif (jusqu'à le multiplier par deux) ; avec des économies de santé dans les ressources allouées aux soins.

Une manière d'opérationnaliser ces données consiste à suivre en continu les évolutions du patient à l'aide d'échelles fiables, valides, et faciles d'utilisation (Cazauvieilh, 2013, 2014, 2015) sur les indicateurs constamment associés aux meilleurs résultats et qui sont donc des prédicteurs des meilleurs résultats (le changement, et l'alliance thérapeutique). Ceci est en lien avec les préconisations de l'APA (Task force, 2006) sur l'excellence clinique et les meilleures pratiques soutenues empiriquement (Evidence Best Practice ou EBP).

1.1.4.2. L'entraînement des thérapeutes par la pratique délibérée (Deliberate Practice, Ericsson, 2006)

Les recherches soulignent la disparité dans l'efficacité des psychothérapeutes et dans la capacité à aider significativement certains patients (niveau micro), tout en suggérant un effet

plateau de l'efficacité des cliniciens, ce niveau d'efficacité serait atteint assez rapidement dans la carrière (niveau macro).

La pratique ROM (SCR) peut être utilisée par le clinicien pour travailler en séance avec son patient et adapter la thérapie afin de viser à obtenir les meilleurs résultats (niveau-micro), mais elle autorise également la mise en place d'un plan d'action de supervision et de formation plus ambitieux et à potentiel de développement de l'efficacité des cliniciens sur le long terme (niveau macro).

Ici également, des données transdisciplinaires relatives au développement de l'expertise issues de domaines de pratiques professionnelles variés, sont susceptibles d'amener des éléments de réponse aux difficultés des cliniciens à développer leur efficacité au cours de leur carrière.

Initialement développée par le psychologue K Anders Ericsson, la notion de « pratique délibérée » est issue de recherches visant à comprendre, à travers des domaines de compétences variés (arts, sports, échecs, musique), comment les praticiens pouvaient acquérir une expertise marquée. Un des premiers résultats issus de ce champ de recherche qui a été popularisée par Malcolm Gladwell dans son ouvrage « Outliers », est que le développement de cette expertise est hautement chronophage, et qu'il faudrait en moyenne 10000 heures de pratique délibérée pour acquérir le statut de virtuose dans un domaine de compétence défini. La pratique délibérée est une forme très spécifique d'entraînement, qui n'est pas définie par la seule notion du temps d'expérience accumulée dans la réalisation d'une tâche (par exemple le nombre d'heures passées à délivrer des séances de psychothérapie), mais plutôt opérationnalisée par le temps dévolu à l'exercice spécifique et délibéré, au seuil de ses capacités, à maîtriser une nouvelle habileté, tout en recevant du feedback adapté (Ericsson, 2006). Cette pratique a été initialement incorporée au champ de la psychothérapie par le Dr Scott Miller (Bertolino, Bargmann & Miller, 2012) en présentant une méthodologie utilisant la pratique ROM pour a) établir une ligne de base des performances individuelles, ce qui revient à repérer des situations ou des

domaines dans lesquels le thérapeute peut se trouver en réelles difficultés au vu d'un critère extérieur objectif de résultats (ce qui va lui permettre ensuite de suivre sa progression), b) s'engager dans des activités de pratique délibérée pour s'entraîner sur une compétence non maîtrisée et c) affiner la séquence d'entraînement en recevant du feedback correctif à chaque étape (à l'aide du suivi de la pratique ROM et/ou d'un mentor/superviseur).

Les premiers résultats de la recherche sur l'amélioration de l'efficacité des cliniciens vont dans le sens d'une association entre la quantité de temps dévolu à la pratique délibérée spécifique et l'efficacité clinique en psychothérapie. Ainsi, il a été démontré chez un groupe de thérapeutes qui monitoraient leurs résultats au sein d'un réseau de recherche basée sur la pratique, que les thérapeutes les plus efficaces passaient plus de temps à développer leurs compétences à l'aide de pratiques spécifiques issues de la pratique délibérée (Chow et al., 2015). Enfin, il est documenté que les performances des cliniciens déclinent avec le temps (Goldberg et al., 2016 a), et une étude plus récente souligne le potentiel de la pratique délibérée pour contrer ce phénomène ; avec un très léger accroissement de l'efficacité des cliniciens au cours du temps relié à la pratique délibérée (Goldberg et al., 2016 b). Par conséquent, cette notion de pratique délibérée est une variable importante à contrôler dans les études sur l'efficacité des thérapeutes, puisqu'elle fait partie des rares déterminants qui semblent liés aux meilleurs résultats et qui sont associés directement aux thérapeutes.

1.1.5. Les sept points clé de la psychothérapie

L'analyse de la littérature relative aux psychothérapies laisse clairement apparaître des données à la fois très positives et assez préoccupantes. En effet la pratique de la psychothérapie peut être associée à quatre résultats extrêmement positifs (ce qui fonctionne) : 1) Suivre une psychothérapie est généralement efficace, aux vues des études rigoureusement contrôlées et pour une problématique spécifique ; 2) La psychothérapie est également généralement efficace dans les conditions réelles où les patients présentent de multiples problématiques ; 3) Une variété de traitement *Bonafide* sont à mêmes d'aider un vaste panel de patients et 4) les acquis

de la thérapie bénéficient aux patients au-delà de leur problématique psychologique et impactent également les coûts de santé.

Cependant, il ressort aussi de cette synthèse que trois éléments négatifs et inquiétants sont associés à la pratique de la psychothérapie : 1) La psychothérapie peut nuire, les effets secondaires persistants existent et les patients sont à risque de se détériorer et d'abandonner leur traitement, 2) les cliniciens se surestiment dans l'évaluation de leurs résultats et présentent de grandes difficultés à discriminer les patients à risque. Enfin, 3) leurs performances ont plutôt tendance à décliner au cours de leur carrière en l'absence d'efforts concertés.

Pour répondre à ces enjeux et reprenant l'intitulé du titre d'un chapitre écrit par Lambert en 2010, nous affirmons qu'il est temps pour les cliniciens d'assurer le suivi des résultats en soins courants, et d'adopter une démarche basée sur le feedback des patients, pour adopter une pratique délibérée aussi bien au niveau micro (à chaque séance) qu'au niveau macro (et travailler à son développement professionnel continu).

L'implémentation des systèmes de suivi des résultats devrait selon nous être encouragée dans les pays francophones, au niveau des cabinets des praticiens, mais aussi au niveau des institutions de soins privées et publiques ; particulièrement en France où les données probantes issues de la pratique de terrain seraient à même d'éclairer les expérimentations courantes quant au remboursement des psychothérapies⁶.

Ces expérimentations permettront peut-être à terme de garantir un accès aux psychothérapies, forme d'intervention qui est, nous l'avons vu, efficace et susceptible de réduire significativement les coûts de santé ; tout en ne rentrant pas dans le champ de la définition trop restrictive du modèle biomédical classique, laissant aux psychothérapeutes leur

⁶ Voir par exemple : <http://www.psychologues-psychologie.net/attachments/article/471/AssuranceMaladie-Guide-pour-psycho-Remboursement-actes.pdf>

expertise et leur autonomie d'exercice. La psychothérapie n'est pas non plus une intervention non médicamenteuse comme les autres, tant la dimension contextuelle et relationnelle est prépondérante.

Enfin, les recherches sur l'efficacité des interventions liées au feedback devraient selon nous être également stimulées, car elles éclairent ce qui fonctionne réellement dans les psychothérapies, au-delà des luttes intestines entre les partisans d'obédiences thérapeutiques différentes, qui sont évidemment stériles pour les premières personnes concernées : les patients. Nous passerons en revue ces recherches dans la section suivante.

1.1.6. Synthèse des données relatives aux fonctionnements et au dysfonctionnements des psychothérapies

- Bien que la psychothérapie soit un traitement qui ait démontré son efficacité par rapport à l'absence de traitement et au placebo, la réponse des patients à leur traitement est très variable, et les taux d'abandon, de détérioration pendant les soins, et les effets secondaires sont loin d'être négligeables.
- L'efficacité des thérapeutes varie considérablement et ne semble pas s'améliorer avec l'expérience, en l'absence d'exercices de pratique délibérée. Les thérapeutes sont soumis à des biais cognitifs quand il s'agit d'identifier les patients à risque de détérioration ou d'évaluer leurs mauvais résultats.
- Deux prédicteurs sont associés aux traitements les plus efficaces : l'alliance thérapeutique entre le patient et son thérapeute qui prédit l'engagement dans les soins ; les résultats positifs précoces au traitement en termes de diminution de la détresse psychologique.
- Le Suivi Continu des Résultats de la thérapie (Routine Outcome Monitoring - ROM) est une méthodologie prometteuse, visant à utiliser le feedback des patients pour recueillir et utiliser l'information apportée par ces deux prédicteurs, et qui est associée à une meilleure efficacité dans les soins.
- Cette méthodologie n'est pas fréquemment utilisée dans les contextes francophones.

Il est en conséquence nécessaire d'approfondir les recherches sur les interventions utilisant le feedback en psychothérapie

1.2. Le Suivi Continu des Résultats en psychothérapie : Revue de la littérature du Routine Outcome Monitoring (ROM) et présentation du système PCOMS.

La pratique du Routine Outcome Monitoring apparaissant, nous l'avons vu, comme une réponse prometteuse face aux enjeux d'évaluation de la psychothérapie dans la pratique clinique, et dans l'amélioration des résultats dus aux traitements, il est maintenant nécessaire de synthétiser l'information sur l'efficacité de l'implémentation de cette méthodologie. Nous allons ainsi dégager les principaux résultats associés à la pratique ROM, dans une perspective de contextualisation des soins.

En effet, nous avons précédemment évoqué que la notion d'efficacité générale de la psychothérapie, la démonstration de la supériorité de la prise en charge face à l'absence de traitement, pour un large panel de problématiques psychologiques, faisait maintenant l'objet d'un consensus (Lambert & Ogles, 2004). Ceci n'exclut pas que les patients puissent répondre en proportions variables aux traitements ou qu'ils soient à risque de se détériorer, voire d'abandonner les soins (Crawford et al., 2016 ; Swift & Greenberg, 2012). Pour répondre à ces enjeux, deux paradigmes ont émergé en parallèle du mouvement orienté vers les traitements empiriquement validés (Empirically Supported Treatment ou EST). En effet, dans l'approche EST, l'objectif est de valider dans des conditions expérimentales rigoureuses, un traitement censé être thérapeutique ; puis de le disséminer dans la pratique clinique courante, en assurant la fidélité au protocole par l'utilisation d'un manuel. Dans cette configuration, c'est la meilleure recherche disponible qui vient « informer » la pratique de terrain. Sans oblitérer l'intérêt des recherches de laboratoire, le mouvement de la recherche centrée sur le patient (Patient Focused Research, Lambert & Hansen, 2001) et le mouvement des preuves basées sur la pratique (Practice Based Evidence, Barkham & Margison, 2008) partagent la particularité d'opérer un retour vers les données de terrain. La pratique clinique courante est alors considérée comme une source inestimable de données probantes souvent négligées. Un des objectifs énoncés de ces approches est en conséquence de refermer le fossé qui sépare la pratique de la recherche en psychothérapie. À cet effet, il s'agit de collecter des données probantes au cours du traitement

par la mise en place d'un système de gestion ou de Suivi en Continu des Résultats (SCR ou Routine Outcome Monitoring - ROM) pour ajuster le traitement. C'est alors la pratique qui vient « informer » les recherches en complétant les meilleures données disponibles issues des pratiques empiriquement validées (EBP).

1.2.1. Historique du Routine Outcome Monitoring

Howard, Moras, Brill, Martinovich, et Lutz ont été les premiers en 1996 à produire un système de suivi des résultats des patients, en s'appuyant sur le modèle dose-effet et sur le modèle en phase de la psychothérapie⁷. En considérant qu'au-delà des preuves d'efficacité des traitements obtenues dans les études contrôlées (Efficacy), et des preuves d'efficacité des pratiques démontrées dans les études naturalistes de terrain (Effectiveness), les cliniciens se doivent de répondre à une question bien plus lourde de sens : « Est-ce-que l'état de ce patient particulier répond (bien) au traitement qui (lui) est appliqué ? ». Il est par exemple documenté que tous les patients ne s'améliorent pas de la même façon, même s'ils reçoivent un traitement *bona-fide* réputé efficace. Pour répondre à cette question, les auteurs ont utilisé les scores obtenus à une échelle de mesure de la santé mentale (le MHI), recueillis auprès d'une importante cohorte de patients suivis en psychothérapie externe afin de modéliser leur trajectoire de changement. Ils ont ainsi pu estimer, sur cette base de données normatives, une réponse probable ou attendue au traitement basée sur les caractéristiques cliniques présentées par les patients au début des soins (l'intercept représentant le score du patient à la première séance et la pente ou slope représentant le taux de changement attendu en fonction du nombre

⁷ Le modèle dose-effet postule une relation log linéaire entre le nombre de séances et l'efficacité au traitement, le modèle en phase postule que les patients traversent trois phases successives au cours de leur traitement (respectivement une phase de remoralisation rapide, de remédiation plus lente et enfin de réhabilitation progressive de leur fonctionnement).

de séances). Générer une trajectoire estimée de traitement pour chaque nouveau patient autorisant ainsi une comparaison entre son état, représenté par son score actuel et la réponse attendue au traitement (calculée à partir de l'échantillon normatif et permettant notamment l'affichage d'un feedback de la progression sous forme d'un graphique ou d'un message écrit) afin de déterminer si le patient nécessite un quelconque changement dans la délivrance des soins. Nous pouvons illustrer cette méthodologie en évoquant l'utilisation des courbes de poids et de taille qui sont reproduites dans les carnets de santé en France, et qui permettent au médecin, de cerner si le patient qui est reçu en consultation, s'inscrit dans la trajectoire d'une croissance normale, ou s'il s'écarte de ce qui est attendu pour sa classe d'âge.

1.2.2. Définition et principes généraux du Suivi Continu des Résultats (ROM)

Le Suivi Continu des Résultats (ROM) est donc une méthodologie qui s'appuie sur la mesure des résultats au cours de la thérapie. Ici les résultats sont (généralement) autoévalués par les patients (de l'anglais Patient Reported Outcome Measures - PROMS ou Résultats de Santé Rapportés par les Patients ; HAS, 2016)⁸. Ces mesures auto-rapportées visent à recueillir l'expérience des patients quant à leur symptomatologie et à leur fonctionnement en lien avec leur qualité de vie, dans la perspective d'évaluer les résultats dus aux traitements. Dans cette approche, le patient est ainsi considéré, à juste titre, comme un expert dans l'expérience de sa propre santé (Kendrick et al., 2016). Il est usuel de parler de feedback pour désigner ce retour d'informations du patient vers le thérapeute et de situer cette approche dans le paradigme des Traitement Orientés par le Feedback que nous pouvons désigner sous l'acronyme TOF (de l'anglais Feedback Informed Treatment ou FIT).

⁸ Pour une présentation des PRO(MS) voire : https://www.has-sante.fr/portail/upload/docs/application/pdf/2016-11/colloquehas_session2_diaporama.pdf

De nombreux systèmes ROM ont été développés, pour être utilisés dans des contextes de soins et de classes d'âges variés. Par exemple, pour les consultations de soins externes, internes, à destination des enfants, des adolescents ou des adultes ; et leur présentation complète dépasserait largement le cadre de cette étude. Néanmoins, le point commun de la plupart des systèmes de suivi (ou de gestion des résultats) est de viser à mesurer les signaux vitaux de la santé mentale et de la thérapie parce qu'il est maintenant clairement démontré que les cliniciens ont beaucoup de difficultés à discriminer ces signaux et à prédire les mauvais résultats de leurs patients (Hannan et al., 2005 ; Lambert, 2017).

1.2.2.1. Les indicateurs à suivre en continu

La myriade de recherches sur l'efficacité des psychothérapies laisse ainsi émerger deux indicateurs que nous avons précédemment évoqués et qui seraient systématiquement associés aux meilleurs résultats à l'issue du traitement : une bonne alliance thérapeutique entre le praticien et son patient, et une réponse initiale positive au traitement.

L'alliance thérapeutique

Il est maintenant établi que l'alliance thérapeutique prédit l'engagement des clients dans les soins, et que l'engagement des clients dans les soins soit un prédicteur de la réussite au traitement (Bertolino, Bargmann & Miller, 2012 ; Orlinsky et al., 2003). L'alliance thérapeutique n'est pas un état fixé mais un processus dynamique en co-construction. Aussi, il est nécessaire de pouvoir évaluer sa qualité, de repérer de potentielles ruptures et d'y faire face.

Le processus de recueil du feedback en temps réel permet de répondre à ces enjeux (Bertolino, Bargmann & Miller, 2012). De plus, l'alliance thérapeutique évaluée par le clinicien prédit avec moins de fiabilité la réussite au traitement que l'alliance évaluée par le patient (Bachelor et Horvath, 1999). Il a été par ailleurs démontré qu'une alliance thérapeutique qui s'améliore en cours de traitement prédit de meilleurs résultats en fin de traitement et moins d'abandon pendant les soins (Duncan et al., 2010). Pour utiliser une analogie de nature

biologique, l'alliance thérapeutique désignerait le « pouls » de la thérapie et sa bonne variabilité serait liée à la bonne santé du processus thérapeutique.

Le degré de détresse psychologique

D'après les travaux de Howard et al. (1986), le changement en psychothérapie suit une loi dose-effet, avec une décélération au fil des séances. Même si les patients ne présentent pas tous une trajectoire identique, il est ainsi attendu que des changements positifs dans la symptomatologie se manifestent dans le cas de traitements efficaces plus tôt que tard ; et que la probabilité d'échec à l'issue du traitement augmente avec la durée (le nombre de séances) pendant laquelle le patient ne rapporte pas de réels changements positifs (Lambert & Ogles, 2004).

En outre, la sévérité de la détresse psychologique mesurée au début du traitement serait un prédicteur plus puissant de l'issue des soins que le diagnostic formel (Bertolino, Bargmann & Miller, 2012). Ainsi, des patients présentant un niveau de détresse plus élevé seraient plus à même de réaliser d'importants changements pendant la thérapie que des patients commençant la psychothérapie avec un faible niveau de détresse (Duncan et al., 2010). Monitorer la détresse psychologique, pourrait permettre de suivre l'évolution réelle des patients et surtout d'ajuster le « dosage », l'espacement des séances dans le temps, de terminer le traitement ou de considérer la nécessité d'impliquer un autre acteur de santé dans les soins. Pour utiliser une illustration naturaliste, le degré de détresse du patient serait lié au potentiel de croissance pendant le traitement et renseignerait sur le dosage des soins à lui prodiguer pour assurer au mieux son développement.

1.2.2.2. Des systèmes de feedback très différents

La plupart des systèmes de gestion des résultats mesurent le feedback du patient sur une symptomatologie liée à la santé mentale, qui renvoie à la mesure d'un construit de détresse psychologique générale. Cette détresse psychologique n'est bien entendu pas ici une constante

fixée (trait) mais plutôt une variable fluctuante (état) qui est censée refléter le changement thérapeutique quand le traitement est efficace. Par exemple, l'Outcome Questionnaire de Lambert ou OQ-45.2 (OQ-45.2, Lambert & Burlingame, 1996a) et le questionnaire d'Impact OQ-22 (Lambert & Burlingame, 1996b ; en français MI-22, Flynn et al., 2003) sont des échelles qui mesurent le fonctionnement d'un patient sur la sphère individuelle (détresse subjective), interpersonnelle et sociale. Et les scores sur ces dimensions sont censés évoluer favorablement si le traitement est efficace pour un individu donné et rester stable pour une personne, qui, devant les mêmes difficultés ne recevrait pas de traitement. L'échelle d'Évaluation des Résultats de la Séance (Outcome Rating Scale ou ORS ; Miller et al., 2003), inspirée du questionnaire OQ-45 de Lambert, mesure un degré général de bien-être à partir de ces trois domaines de fonctionnement. Quand ce degré de bien-être est mesuré de telle sorte qu'un plus haut score au questionnaire renvoie à un meilleur bien-être perçu, le score est également censé se situer au-dessus du seuil clinique et rester relativement stable chez des patients ne présentant pas de problématique marquée ; alors qu'il est censé être significativement plus bas que le seuil clinique chez les patients suivis en thérapie, et il est attendu qu'il évolue à la hausse en cas de réussite due au traitement.

Certains systèmes de gestion des résultats comprennent des échelles complémentaires comme par exemple une mesure de l'alliance thérapeutique. Le système PCOMS (Partners for Change Outcome Management System ou Partenaire pour le Changement Système de Gestion des Résultats, Miller et al., 2005) comprend à la fois un questionnaire d'évaluation des résultats de la thérapie et un questionnaire de mesure de l'alliance (respectivement les questionnaires ORS et SRS ; Miller et al., 2005). Ce système permet de recueillir et d'utiliser le feedback des patients pendant les soins pour mesurer, de séance en séance, le changement psychothérapeutique (augmentation, du degré de bien-être mesuré par l'échelle ORS) et l'évolution du rapport collaboratif (avec l'évolution positive des scores mesurés par l'échelle d'Évaluation de la Séance, ou Session Rating Scale ; SRS, Miller et al., 2005). Le processus

clinique impliqué dans l'utilisation du feedback est alors quelque peu différent puisque sont utilisées par le praticien à la fois les données de suivi du changement, par exemple à chaque début de séance (avec l'échelle ORS), et à chaque fin de séance (avec l'échelle SRS) les données sur la qualité de la relation thérapeutique. Ces données sont par ailleurs aussi bien renvoyées au thérapeute qu'au patient. Ceci permet notamment d'engager le patient dans les soins et de créer, renforcer et de réparer la relation thérapeutique. Dans le cas du système OQ-45, des outils complémentaires de support clinique (Clinical Support Tools) permettent également d'évaluer l'alliance thérapeutique, les besoins, et la qualité du réseau social, mais sont eux administrés uniquement si le patient s'éloigne significativement de la trajectoire espérée de changement fiable prédite par l'algorithme.

Le feedback recueilli peut donc également différer sensiblement selon les dispositifs, en étant « renvoyé » uniquement au clinicien, au patient, ou aux deux parties prenantes (comme dans le système PCOMS) ; avec l'implication de processus de changement associés différents. De même, tous les systèmes de feedback ne se basent pas sur la même temporalité (par exemple avec un recueil au début et à la fin de la séance, un feedback immédiat ou renvoyé avec un délai entre les séances).

En outre, les mesures peuvent être administrées sous la forme traditionnelle de questionnaires papier-crayon et les résultats sont analysés par le clinicien, en fonction de seuils et d'indices de changement fiables fournis par les concepteurs des questionnaires ; et établis à partir de données normatives sur des échantillons cliniques et non cliniques (et notamment calculés à partir des formules de Jacobson, 1991). Ou bien administrées à l'aide d'outils informatiques (ordinateur, tablette, smartphone). Et, ici, en fonction des données normatives disponibles, la trajectoire actuelle du patient peut être comparée aux données obtenues à partir de larges échantillons de patients ayant suivi une thérapie (benchmarking). Dans ces conditions, le logiciel peut même utiliser un algorithme prédictif et alerter les cliniciens sur l'éventualité d'un mauvais résultat à l'issue du traitement, désignant les patients à risque de mauvais résultats

ou NOT (de l'anglais NOT ON TRACK = PAS EN BONNE VOIE de rétablissement) ou ceux qui répondent positivement au traitement ou ON (de l'anglais ON TRACK = EN BONNE VOIE de rétablissement) ; et/ou suggérer une marche à suivre pour remédier à la situation (par exemple discuter des objectifs du traitement avec le patient ou considérer une consultation médicale).

Enfin, selon les dispositifs techniques, il est possible d'envisager le calcul de l'efficacité clinique, aussi bien au niveau individuel que pour un groupe de cliniciens ou une institution de santé, ce qui permet de documenter les effets dus au traitement ; ou de travailler en continu selon les principes de la pratique délibérée pour acquérir de nouvelles compétences (Ericsson, 2006). Les systèmes de feedback et les contextes d'utilisation varient donc considérablement ; aussi il serait hasardeux d'interpréter les résultats liés à l'efficacité de ces interventions sans tenir compte de ces variables potentiellement confondantes. Enfin, il est à noter que les systèmes de feedback ont initialement été développés pour les patients suivis en consultation externe (counseling ou psychothérapie).

1.2.3. Revue générale de la littérature sur les effets liés aux pratique ROM

Poser des questions différemment amène à plus ou moins restreindre le champ des réponses. Ainsi, interroger un phénomène par une question fermée (par exemple « est-ce que cette intervention est efficace » ?) ne permet pas d'envisager une réponse qualitativement graduée (« efficace selon quels contextes ? »). Aussi, il est nécessaire d'être prudent, en considérant les différents dispositifs d'analyse des effets dus au feedback, et avant de généraliser les résultats. Les essais contrôlés, randomisés ou non, et les méta-analyses de ces mêmes essais aboutissent à des traitements de l'information très différents. Ainsi, agréger les résultats des ECR, dans une méta-analyse sur les effets du feedback, avec par exemple des patients suivis en consultation interne et externe peut aboutir à une taille de l'effet réduite et non significative si l'intervention est efficace dans un contexte de soin, mais pas dans l'autre ; et cela souligne l'intérêt de l'analyse des modérateurs. La conclusion générale à retirer de ce

résultat peut cacher un autre résultat intéressant et cliniquement significatif. Aussi, le choix des critères d'inclusion dans les revues et les méta-analyses est prépondérant (le fameux « garbage in- garbage out », littéralement « donnés inutiles en entrée, données inutiles à la sortie ») ; tout comme la vérification de la comparabilité des résultats (« le système de feedback est-il structurellement implémenté de la même façon - utilisé de manière similaire dans ces contextes de soins différents ? »).

Pour éclairer quelque peu ces points, nous allons croiser les données issues des revues de la littérature concernant la pratique ROM (incluant plus spécifiquement l'OQ-45 et le système PCOMS), en évoquant la recension de la littérature de Schuckard & Miller (2016), et les méta-analyses disponibles (Knaup et al., 2009 ; Lambert et Shimokawa, 2011 ; La revue systématique Cochrane et la méta-analyse de Kendrick et al., 2016 ; Østergård et al., 2018).

Il n'est pas, selon nous, souhaitable de présenter ici toutes les études répertoriées, ce qui reviendrait à reproduire le travail de revue systématique et/ou de méta-analyse à partir d'éléments déjà plusieurs fois « revus » et « méta-analysés » ; mais nous pouvons réaliser une analyse comparative critique des différences de résultats, contextuellement circonstanciée et à même d'éclairer les variables les plus significatives quant à l'utilisation du SCR en général et du système PCOMS en particulier.

En premier lieu, la sélection des études est bien entendue importante et les critères divergent selon les auteurs. Par exemple, la décision d'inclure une étude sur la base de la présence d'une mesure PROM ou dite de « feedback » est parfois hasardeuse. Ainsi, plusieurs études issues de la méta-analyse de Knaup et al. (2009) par exemple, posent problème selon Schuckard & Miller (2016), parce qu'il ne s'agirait pas d'études relatives au feedback. Par exemple, l'article d'Ashaye, Livingston & Orrell (2003) inclus dans cette méta-analyse, documente un ECR randomisé chez des patients âgés de 65 ans et plus, suivis en hôpital de jour ; la mesure utilisée est une échelle d'évaluation des besoins des personnes âgées qui est en fait cotée par le clinicien (hétéro-évaluée) ; l'évaluation ne répond pas à l'acceptation générale

des mesures auto-rapportées, ou de feedback, désignant les PROMS. Bien que de manière quelque peu surprenante, dans la revue Cochrane, les PROMS soient définies principalement comme des mesures évaluées côté patient mais qui peuvent également faire l'objet d'une évaluation et d'une interprétation par le thérapeute, avec le patient. La définition commune des PROMS et du National Health Service anglais (NHS), par exemple, est plus restrictive puisqu'elle désigne les échelles auto-rapportées (Kingsley & Patel, 2017). L'étude de Marshall et al. (2004), également incluse dans cette méta-analyse, est critiquée comme n'étant pas, non plus une étude de feedback par Schuckard & Miller (2016) : dans cette recherche le « feedback » est une évaluation standardisée collectée et générée par une infirmière au coordinateur de soins. L'étude de Brodey et al. (2005) qui est pourtant reprise dans la méta-analyse de Knaup et al. (2009) est disqualifiée par la revue Cochrane (Kendrick et al., 2016).

Ensuite, les études peuvent être exclues car elles correspondent plutôt à de la littérature « grise » même si elles sont accessibles (non publiées dans une revue à comité de lecture par exemple). Ainsi, l'étude de Bauer et al. (2004), avec des patients âgés de 18 à 79 ans suivis en consultation interne, est également critiquée par Schuckard & Miller (2016) car c'est une étude non publiée (dissertation de thèse en allemand). Selon Knaup et al. (2009), qui l'incluent dans leur analyse, elle présenterait une taille de l'effet due au feedback non significative et très faible sur la santé mentale à court terme (phase 3 et 4) et une taille de l'effet due au feedback faible et positive sur la durée du traitement (en phase 4). L'étude de Trudeau (2001) avec 127 patients adultes suivis en consultation externe est également une dissertation non publiée (avec des effets non significatifs) et critiquée pour cette raison. Nous pouvons noter, enfin, une très grande hétérogénéité d'échelles de mesures considérées dans certaines études comme dans la méta-analyse de Knaup et al. (2009) ; échelles de besoins, de symptômes, échelles d'alliance, de mesures du support social, etc) de même à affecter les résultats finaux puisque les instruments de mesure (échelles différents) peuvent amener à des tailles de l'effet différentes. Certaines

études peuvent être également exclues parce qu'elles ne répondent pas à la définition formelle de l'ECR.

1.2.3.1. Principaux résultats de la littérature

La recension de Schuckard & Miller (2016) sur la pratique ROM et le système PCOMS aboutit à une synthèse clairement positive, avec dans l'ensemble des effets prometteurs dus à l'utilisation des systèmes ROM (Miller, et al., 2015). Ainsi, ils suggèrent que, selon les études, l'utilisation du Routine Outcome Monitoring serait associée à une diminution des taux d'abandon (qui pourrait atteindre 50%), à une diminution du nombre de patients qui voient leur état se détériorer pendant les soins (jusqu'à 33 %) et à une augmentation du taux de changement cliniquement significatif qui pourrait être doublé (défini par la conjonction du passage du statut clinique à non clinique et d'un changement dépassant un score de changement fiable). Des gains secondaires sont également rapportés au regard des dépenses de santé qui seraient diminuées (Cazauvieilh, 2013, 2014, 2015).

1.2.3.2. Des effets positifs pour les patients NOT (PAS EN BONNE VOIE de rétablissement)

Des effets positifs liés aux patients qui ne sont pas en bonne voie de rétablissement, c'est-à-dire qui ne s'amélioraient initialement pas au début de la thérapie, sont notamment rapportés dans la méta-analyse de Lambert et al. (2003) qui reprend trois ECR, compilant les données obtenues avec une population de 2602 patients suivis en consultation externe pour des difficultés personnelles, avec le système OQ-45 (feedback au clinicien). Cette méta-analyse présente un résultat positif avec une taille de l'effet moyenne de .39 en faveur de la condition feedback, c'est-à-dire quand les thérapeutes obtiennent le signal (négatif) que les patients ne sont pas en train de s'améliorer (en comparaison de la condition non feedback). Le taux de détérioration est également réduit (de 8%) quand sont comparées les conditions feedback et non feedback pour ces patients NOT. Il est à signaler que, dans une des études (Whipple et al.,

2003), le feedback est complété en cas de mauvaise réponse du patient à la thérapie par une aide clinique supplémentaire apportée aux cliniciens (Clinical Support Tools), sur l'alliance thérapeutique, les besoins du patient et son support social. Enfin, dans cette étude, le feedback exercerait également un effet sur les résultats déjà positifs au traitement puisqu'aussi bien les patients NOT que ON, c'est-à-dire l'échantillon total en condition feedback, seraient améliorés par rapport au groupe contrôle (avec une taille d'effet faible mais significative de 0.09 en faveur du feedback).

1.2.3.3. Des effets positifs pour tous les patients (NOT et ON)

Des effets positifs dus au feedback pour tous les patients (NOT et ON) sont également retrouvés dans l'étude de Berking et al. (2006), avec une population clinique souffrant d'anxiété et de dépression suivie en soins externes. Ici la taille de l'effet est importante (de .48 à .50 selon les questionnaires utilisés) ; et un effet positif du feedback sur la santé mentale à court terme est également rapporté ($g = .45$, Intervalle de Confiance : $IC (.06, .85)$, $p = .02$, Knaup et al., 2009). Le même type d'effet sur la population totale est rapporté dans l'étude de Harmon et al. (2007), pour des patients suivis en soins externes et présentant des problématiques variées (trouble de l'humeur, anxiété, troubles de l'ajustement). Le feedback enrichi par les outils de support clinique est ici efficace et double le taux de changement cliniquement significatif. Mais la provision au patient du feedback, en plus de la provision au clinicien, n'a pas d'effet significatif. Il n'y a au final pas d'effet du feedback sur le nombre de séance chez les patients en condition ON TRACK » mais un effet constaté avec les patients NOT. L'effet faible du feedback sur tous les patients, avec une taille de l'effet de .10, est également retrouvé dans la méga-analyse de Lambert & Shimokawa (2011), comprenant les données de 6 ECR ; avec un effet fort pour les patients NOT (de .40 à .70 selon Schuckard & Miller, 2016).

1.2.3.4. Des effets positifs quand le feedback est renvoyé au patient

Des effets positifs modérément accentués quand le feedback est renvoyé aussi bien au thérapeute qu'au patient, sont retrouvés dans l'ECR de Hawkins et al. (2004) avec 201 patients adultes suivis à l'hôpital (interne) pour des troubles de l'humeur ($d = .33$). Dans cette recherche, les progrès sont mesurés par l'OQ-45 et des recommandations sont renvoyées au thérapeute (graphiques et textes), tandis que des messages écrits spécifiques sont renvoyés aux patients. L'étude documente aussi une amélioration globale pour tous les patients en condition feedback ($d = .31$). Par contre il n'y a pas d'effets significatifs constatés sur la santé mentale à court terme ($g = .14$, IC (-.20, .48), $p = .43$; Knaup et al., 2009). À signaler que la durée du traitement des patients NOT est doublée du fait du feedback.

1.2.3.5. Une réduction des dépenses de santé

Des effets positifs liés aux coûts de santé sont notamment retrouvés dans l'étude de Slade et al. (2006) avec une population de patients suivie en soins externes et présentant des troubles psychotiques. Le feedback graphique et textuel des besoins des patients, de leur qualité de vie et de l'alliance est renvoyé au clinicien. L'intervention n'entraîne pas de différence significative dans les résultats mesurés par trois questionnaires mais elle réduit les admissions et la durée de séjour (de 16,4 jours à 3,5 jours), avec une réduction du coût en condition feedback rendant l'intervention efficace. Il n'y a pas eu d'effets significatifs sur la santé mentale à long terme constatés par ailleurs ($g = -.03$, IC (-.37, .32), $p = .88$, Knaup et al. (2009).

1.2.3.6. Un effet sur la symptomatologie spécifique

Des améliorations dans la rémission de certains symptômes en condition feedback sont également rapportées ; par exemple dans l'étude de Schmidt et al. (2006) sur la restriction alimentaire et les vomissements, bien que l'effet global sur la santé mentale à court terme ne soit pas significatif ($g = .37$, IC (-.31, 1.06), $p = .28$), Knaup et al., 2009). Dans l'étude de Brodey et al. (2005), les symptômes dépressifs sont réduits, avec un effet faible mais

cliniquement significatif. Il est à noter que cette étude a été écartée de la revue Cochrane mais qu'elle est listée par Knaup et al., 2009 (qui rapportent par ailleurs un effet non significatif du feedback sur la santé mentale à court terme).

1.2.3.7. Un effet positif du au feedback instantané

Des effets supplémentaires sont constatés quand le feedback est renvoyé de manière immédiate à l'aide d'un outil informatique au clinicien, plutôt que de manière espacée dans le temps pour les patients en condition ON (Slade et al., 2008).

1.2.3.8. Un effet global sur le traitement

Enfin, des effets positifs compilés sur les résultats et la diminution des durées de traitements sont retrouvés dans la méta-analyse de Lambert & Shimokawa (2011) de 9 ECR pour les systèmes OQ-45 et PCOMS (Miller et al., 2015).

1.2.3.9. La méta-analyse de Knaup et al. (2009)

La méta-analyse de Knaup et al. (2009) compile ainsi les données de 12 ECR (patients suivis en interne et en externe, avec une grande variété de présentation dans la sévérité des troubles psychologiques) et rapporte une taille d'effet faible mais significative sur la santé mentale mesurée à court terme ($d = .10$) ; avec un effet non significatif sur la santé mentale à long terme et un effet non significatif sur la durée du traitement. Par contre, il est suggéré dans l'analyse des modérateurs que le feedback apparaîtrait comme plus efficace sur la santé mentale à court terme, quand il comprendrait des informations sur les progrès du patient, et non pas seulement des informations de simple renvoi de son statut clinique (score) ; quand les informations seraient renvoyées aussi bien aux patients qu'au thérapeute, et de façon continue (plus d'une fois).

1.2.3.10. La revue et la méta-analyse Cochrane (Kendrick et al., 2016)

La revue Cochrane (Kendrick et al., 2016) compile les données de résultats de 17 études et 8787 participants présentant des troubles mentaux (les plus communément représentés) et une méta-analyse de 12 ECR incluant 3696 participants utilisant l'OQ-45 et/ou le système PCOMS. Le résultat principal de l'analyse agrégée laisse apparaître qu'il n'y a pas de différence significative de résultats, en termes de symptomatologie, entre les conditions feedback et non feedback (*différence moyenne standardisée* = -0.07, 95% IC (-0.16 - 0.01) ; $p = 0.10$), sachant que l'information disponible pour émettre ce résultat est jugée par les évaluateurs comme étant de faible qualité ; en raison du risque de biais dû à l'absence de dispositif de contrôle en aveugle pour les praticiens et de la perte de données en cours d'étude (attrition). En outre, une seconde méta-analyse de sept études investiguant l'impact du feedback dans la gestion des soins, et incluant 2608 patients, ne révèle pas de différence significative due au feedback (*différence moyenne* = -0.02 sessions, 95% IC (-0.42, 0.39) ; $p = 0.93$), avec des données jugées de faibles qualités. Enfin, avec une sous analyse post-hoc incluant les données de 923 participants, il est retrouvé un résultat qui va dans le sens d'un consensus quant à l'efficacité des systèmes ROM : les résultats seraient significativement moins bons pour les patients à risque de détérioration en début du traitement (ou en condition NOT) et qui ne bénéficient pas de la condition feedback (dans 10 études utilisant soit l'OQ-45, soit le PCOMS et avec une moyenne standardisée des résultats de -0.22, 95% IC (-0.35 , -0.09 ; $p < .01$). Indice d'hétérogénéité $I^2 = 0\%$) ; les données disponibles sont également jugées de basse qualité.

Du fait de ces limitations, les auteurs concluent qu'ils disposent de preuves insuffisantes pour recommander l'utilisation du SCR dans le traitement des troubles mentaux communs ou pour la gestion des soins, mais que ces résultats sont quand même sujets à une incertitude considérable en regard du risque très élevé de biais. Et que de nouvelles études, qui pourraient de ce fait avoir un impact important, sont donc nécessaires. Ils recommandent de résoudre les problèmes liés aux pertes de données, de documenter les effets secondaires dans les recherches, les coûts et de suivre les patients plus longtemps (plus de 6 mois dans le temps). À ce sujet,

l'étude de Chang et al. (2011) documente avec une qualité modérée des données, qu'il n'y a pas d'augmentation du risque suicidaire dans une cohorte de sujets en condition feedback. Ce point va dans les sens de la revue de Dazzi et al. (2014), sur l'impact de l'auto-évaluation des patients, documentant qu'interroger les patients sur leur propre santé mentale n'aurait pas d'impact négatif (et pourrait même avoir des conséquences positives).

Les auteurs tentent cependant d'expliquer ces résultats moins positifs dans les effets dus au feedback par plusieurs facteurs potentiels : des études différentes ont été incluses depuis les méta-analyses initiales ; des résultats divergents sont dus aux critères d'inclusion divergeant de ceux utilisés pour les méta-analyses initiales ; avec par exemple, des études qui ne rentrent pas dans leur définition formelle des ECR et qui présenteraient donc un risque plus élevé de biais (par exemple dans la méta-analyse de Lambert & Shimokawa (2011), chez Harmon et al. (2007), Slade et al. (2008)). En effet, dans certaines études, le groupe contrôle était constitué de données de patients archivées, et non pas de patients avec une randomisation formelle. Enfin, une troisième explication serait liée aux effets d'allégeance des auteurs des échelles à leur propre rationnel de traitement qui pourrait donc influencer les résultats (sans pouvoir toutefois retrouver de véritables différences dans les résultats entre les interventions données par les concepteurs de ces systèmes de feedback qui irait dans le sens de cette hypothèse).

Ces données vont, en outre, dans le sens d'études récentes concernant l'utilisation du feedback en contextes de soins psychothérapeutiques, où le SCR bénéficiait uniquement aux patients NOT (par exemple Davidson, 2015). Il est rappelé que certaines caractéristiques sont considérées comme des modérateurs potentiels de l'effet dû au feedback (De Jong, 2012) : des caractéristiques liées aux thérapeutes, sont en effet associées à l'amélioration des patients en condition NOT (l'engagement à utiliser le feedback ou feedback propensity chez le thérapeute, la validité perçue du feedback, et l'auto-efficacité des thérapeutes). Enfin ils citent les suggestions de Krägeloh (2015) quant au fait que le contexte, par exemple le contexte de soins, ou la structure, permette une plus ou moins grande latitude de discussion du feedback avec les

patients ; l'utilisation de supports informatiques, la fréquence du feedback, ou la discussion des PROMS avec les cliniciens seraient à même d'amener à des résultats différents. À ce sujet, il n'a pas été retrouvé de différences entre les résultats des thérapeutes en condition enrichie par les outils de support clinique (CST) ou non et pas de différences de résultats entre les conditions où le feedback serait donné à la fois au patient et au clinicien. Cela n'exclut pas le rôle potentiel de la flexibilité du thérapeute dans la distribution et l'utilisation du feedback et nous allons éclairer ces éléments de contextualisation des interventions, à même de modérer leur efficacité, par les dernières données disponibles quant au système PCOMS.

1.2.3.11. La méta-analyse de Østergård et al. (2018)

Il s'agit d'une revue systématique de la littérature suivie d'une méta-analyse des résultats du système PCOMS (ORS et SRS), comprenant les données de 14 ECR et de 4 études contrôlées non-randomisées. La conclusion générale de l'étude est que l'utilisation du feedback PCOMS améliore significativement les résultats, en termes de symptômes généraux ($g = .27, p = .01$), par contre, dans cette étude il n'y aurait pas d'amélioration significative des clients NOT, due au feedback, avec une importante hétérogénéité entre les études. L'analyse de modération révèle un effet non significatif de PCOMS en contexte psychiatrique (même avec une allégeance non négative des chercheurs à PCOMS ; $g = 0.10, p = .144$), et un effet positif du feedback en contexte de counseling ($g = 0.45, p < .001$), qui pourrait être lié à des effets d'allégeance ou à l'utilisation unique de PCOMS comme mesure des résultats. Les auteurs suggèrent plusieurs explications à ces résultats : les patients suivis en psychiatrie pourraient présenter des troubles plus sévères et nécessiter des soins plus directifs alors que PCOMS encourage un rapport collaboratif. Il pourrait aussi être démoralisant pour ces patients de recevoir un feedback négatif sur leur progression, du fait de leur état plus altéré. Et enfin, est avancée une hypothèse qui nous semble de première importance, à savoir que le contexte de soins psychiatriques, avec des facteurs organisationnels liés notamment au (manque de) temps pourrait compliquer la nécessaire flexibilisation du traitement ; les auteurs citent Davidsen et

al. (2017) qui suggèrent qu'un programme standardisé de soins pourrait empêcher le thérapeute d'utiliser le feedback pour modifier le traitement (ce qui selon nous constitue l'essence de ces interventions). En outre, dans ce contexte de soins pluridisciplinaires avec l'interaction des patients et d'acteurs très différents, les effets de PCOMS pourraient être « brouillés ».

Synthèse des données d'efficacité des systèmes ROM

A l'exception de la méta-analyse précédente, un consensus semble donc se dégager quant à l'impact significatif des systèmes de feedback sur les résultats des traitements psychothérapeutiques pour les patients NOT (c'est à dire qui ne s'améliore initialement pas au début de leur traitement). Il est à noter que ces patients risquent d'abandonner le traitement ou peuvent se détériorer pendant les soins.

Même si les données disponibles ne convergent pas franchement et que la qualité variable des études ne permette pas de trancher avec certitude, le Suivi Continu des Résultats apparaîtrait plus effectif quand il implique aussi bien le thérapeute que le patient, avec une administration au cours de la séance. Il apparaîtrait également plus efficace s'il était renvoyé en continu au cours du traitement et comprendrait des informations sur les progrès des patients, en plus des données de scores bruts, qui seraient éventuellement partageable avec ces derniers. Ceci souligne qu'il ne suffit pas d'administrer des échelles de feedback pour obtenir de meilleurs résultats mais que de nouvelles recherches sont nécessaires en vue de découvrir quels sont les contextes qui favorisent les effets du feedback en thérapie (les modérateurs). En effet, la littérature n'apporte pas non plus de données convaincantes sur le fonctionnement de cet effet feedback en psychothérapie, et de nouveaux modèles d'analyse du feedback en thérapie pourrait ainsi être proposés à partir des données que nous avons recensées, et être potentiellement éprouvés.

Les auteurs de la revue systématique Cochrane concluent également que de nouvelles recherches incluant des patients, notamment traités avec des médicaments, comme c'est souvent le cas dans la prise en charge des troubles mentaux communs en première intention,

doivent être développées dans les contextes de soins primaires, et suggèrent qu'il est préférable d'utiliser des PROMS plus court, telle que l'échelle ORS, qui peuvent être complétées rapidement par les patients et dont le feedback pourrait être utilisé par les cliniciens, dans les contextes de soins où le temps et les ressources sont limités.

En effet, l'échelle ORS, qui est partie intégrante du système PCOMS inscrit au répertoire américain des pratiques et des programmes empiriquement validés, possède des qualités psychométriques documentées, et est accompagnée d'un programme d'implémentation associé à des données relatives aux coûts d'utilisation et de formation⁹. Ce système a par ailleurs démontré son utilité dans les contextes de soins où les échelles plus longues risquent d'être abandonnées par les thérapeutes (Brown, Dreis, & Nace, 1999 ; Miller et al., 2003 ; Seidel et al., 2016). Le système PCOMS a notamment été validé en Allemand (Janse et al., 2014), et plus récemment en espagnol (Moggia et al., 2018). Ces outils ouvrent la voie à de nouvelles utilisations et recherches sur la pratique ROM dans des contextes de soins variés, à même d'amener à une meilleure compréhension des effets du feedback en psychothérapie.

Nous allons dans la section suivante passer en revue les différentes tentatives de modélisation des effets dus au feedback dans la pratique ROM et tenter de proposer un nouveau modèle explicatif.

⁹ National Registry of Evidence-Based Practice and Programs (EBP-NREPP)
<http://www.nrepp.samhsa.gov/ViewIntervention.aspx?id=249>

1.2.4. Synthèse des données relatives à l'efficacité du Suivi Continu des Résultats en Psychothérapie (ROM)

- L'utilisation des systèmes de feedback est principalement associée à de meilleurs résultats pour les patients qui étaient initialement à risque de se détériorer au cours de la psychothérapie ; avec des gains en efficacité modérés. Ce résultat n'est pas répliqué dans toutes les recherches.
- Les contextes de soins semblent influencer l'efficacité de ces interventions avec de meilleurs résultats obtenus en consultations externes.
- L'efficacité liée au feedback est potentiellement modérée par la fréquence d'utilisation (continue), des informations supplémentaires susceptibles de guider le clinicien dans l'ajustement du traitement, et éventuellement par le partage des informations avec le patient.
- La notion de coût efficacité est primordiale dans la décision d'implémentation des systèmes de feedback, et les systèmes de feedback demandant peu de temps à être administrés comme le système PCOMS, devraient être favorisés.

Il apparaît ainsi nécessaire de développer de nouveaux protocoles rigoureux d'analyses des effets dus au feedback pour améliorer son efficacité.

1.3. Vers un modèle explicatif de l'effet du Routine Outcome Monitoring : De la communication non verbale au feedback instrumental.

1.3.1. Données relatives à l'efficacité du Routine Outcome Monitoring dans les traitements orientés par le feedback

De nouvelles recherches sont ainsi nécessaires pour développer un modèle explicatif des effets dus au feedback dans la pratique ROM, notamment pour comprendre pourquoi le feedback n'améliorerait que les patients à risque de se détériorer. Il n'existe pas non plus, à notre connaissance de modèle générique rendant compte de ces résultats. Nous allons ainsi dans cette section, analyser les variables potentiellement liées à l'efficacité des interventions ROM et proposer un modèle générique inspiré de données transdisciplinaires.

Un nouveau paradigme a émergé des données de la littérature suggérant qu'une proportion significative de patients ne s'améliore pas, se détériore ou abandonne les traitements psychothérapeutiques (Cooper, 2008), dont la finalité était d'identifier les patients à risque de mauvais résultats en cours de traitement, en utilisant leur feedback pour faciliter la prise de décision. La pratique du Routine Outcome Monitoring implique donc le Suivi Continu des Résultats (SCR) de la psychothérapie en utilisant des données fournies par le patient. Même si l'implémentation de ces pratiques amène à des résultats quelque peu divergents, nous l'avons vu, un consensus semble se dégager quant à l'intérêt de ces méthodes pour réellement améliorer les résultats des soins des patients qui ne tiraient initialement pas de bénéfices de leur traitement (Kendrick et al., 2016). Ces patients sont souvent qualifiés dans les recherches de « NOT ON TRACK » ou « NOT » pour désigner qu'ils sont « en mauvaise voie » de guérison ou de rétablissement.

1.3.1.1. Les caractéristiques reliées au thérapeute dans l'utilisation du feedback

Tous les cliniciens ne bénéficient pas de l'apport dû au feedback en proportions égales, et il semblerait que certaines variables liées aux thérapeutes modèrent les effets et l'utilisation du feedback en psychothérapie ; de meilleurs résultats sont par exemple constatés dans les

traitements orientés par le feedback délivrés par les femmes, chez les praticiens qui possèdent une bonne auto-efficacité, et qui ont une meilleure propension à recevoir (accepter) du feedback externe (ou External Feedback Propensity ; De Jong et al., 2012). En particulier, les cliniciens qui auraient plus d'inclinaison à considérer le feedback reçu de l'environnement plutôt que de s'appuyer sur leur propre feedback interne (Internal Feedback Propensity) solliciteraient plus de feedback et seraient plus performants sur de nouvelles tâches (Herold & Fedor, 2003).

1.3.1.2. Les thérapeutes n'apprennent pas du feedback

En outre, bien que les thérapeutes aient la plus grande difficulté à détecter les patients à risque de se détériorer pendant le traitement (Hatfield et al., 2010), ce qui remet en cause l'idée que le jugement clinique soit suffisant, ils ne semblent pas apprendre de leur exposition aux systèmes de feedback. Dans les études contrôlées pour lesquelles les thérapeutes prodiguent leurs soins, en étant à la fois en condition feedback instrumental (recevant un signal d'alarme automatisé en cas de mauvais résultats attendus) pour 50% de leur patientèle et en condition non-feedback (suivant simplement leurs patients de manière usuelle) pour l'autre moitié, les thérapeutes obtiennent des résultats significativement supérieurs dans la condition feedback ; et ce même s'ils sont exposés à cette technologie sur une durée conséquente de plusieurs années (Lambert, 2017). Les thérapeutes ne deviendraient donc pas plus efficaces dans la détection des patients à risque du fait de leur utilisation du feedback. Le feedback fonctionnerait ainsi à la manière d'un système de navigation GPS, qui peut suggérer un meilleur itinéraire en cas de déviation de la trajectoire, mais n'affecte en rien les capacités de navigation générale, ou les capacités à mieux appréhender le territoire de celui qui le reçoit ; et le feedback serait par ailleurs insuffisant dans l'amélioration des performances globales des thérapeutes puisqu'il serait nécessaire de mettre en place, en plus du ROM, un plan d'entraînement spécifique et ciblé dénommé « pratique délibérée » pour améliorer les résultats (Miller et al., 2013 ; Goldberg et al., 2016).

1.3.1.3. Le feedback améliore parfois les résultats sans autre entraînement formel supplémentaire

Enfin, et de manière plus surprenante, au vu des données précédentes, les thérapeutes initialement les moins efficaces dans certaines études contrôlées randomisées, et accédant au feedback instrumental de leurs patients, seraient capable d'atteindre l'efficacité de leurs collègues naturellement plus efficaces, sans autre entraînement formel supplémentaire que la capacité qui leur a été donnée d'utiliser ce système de rétroaction (Anker, Duncan & Sparks 2009, Reese et al., 2010). Ainsi dans l'étude contrôlée randomisée de thérapies de couple d'Anker, Duncan & Sparks (2009), les données d'un échantillon réduit de dix thérapeutes ont suggéré que les praticiens qui démontraient l'efficacité initiale la plus basse bénéficiaient significativement plus du feedback que leurs collègues naturellement plus efficaces, et qu'ils pouvaient atteindre grâce à ce système une efficacité multipliée par quatre (en termes de changement cliniquement significatif en comparaison de la condition non-feedback). Dans l'étude de réplique de Reese et al. (2010), le gain dû au traitement était doublé dans la condition feedback par rapport à la condition non feedback, et ce gain en efficacité a été observé avec des étudiants qui avaient reçu 8 fois moins d'entraînement à l'utilisation du feedback, que les thérapeutes de l'étude d'Anker, Duncan & Sparks (2009) ; toutefois, il est à noter qu'ils bénéficiaient d'une supervision centrée sur l'utilisation de ce système de feedback (PCOMS). Ces résultats sont de nature à suggérer qu'un entraînement formel n'est pas forcément nécessaire pour devenir un thérapeute plus efficace, et qu'un moyen de mieux percevoir des informations pertinentes, ou de les utiliser pour ajuster le traitement peut suffire à augmenter l'efficacité. Ceci rejoignant le point de vue de Worthen & Lambert (2007) sur le fait que le feedback des clients influencerait les résultats en apportant aux cliniciens des informations qu'ils pourraient avoir initialement ignorées ou auxquelles ils n'auraient pas accordé une juste valeur.

1.3.2. Comment intégrer ces données et développer un modèle rendant compte de l'effet du feedback dans la ROM et dans la psychothérapie en général ?

Afin de proposer de nouveaux éléments de réponse à ces questions, nous allons commencer par passer en revue les différents apports de la littérature aux Traitements Orientés par le Feedback (TOF).

1.3.2.1. Modèles théoriques des interventions liés au feedback

Les systèmes de rétroaction sont des processus régulateurs très importants chez les êtres vivants et participants au bon fonctionnement des systèmes dynamiques (biologiques). Dans cette optique, la définition et l'opérationnalisation du concept de feedback par Carver et al. (1981, 1982), nous semble très adaptée au domaine de la psychothérapie. En effet, dans la théorie de l'auto-régulation et du contrôle (Self Regulation and Control Theory), le feedback est simplement considéré comme un mécanisme qui motiverait les individus à changer leur comportement (actuel) ; et le parallèle peut aisément être réalisé avec l'utilisation des systèmes de ROM qui visent à avertir, motiver et aider le clinicien à adapter son traitement. Le processus de changement par le feedback est donc ici spécifiquement un traitement avant tout de nature motivationnelle. Pour expliciter plus avant cette définition et en préciser la dimension de traitement de l'information (cognitive), ce serait l'existence d'un écart significatif entre un but considéré comme important et la situation présente qui induirait un état de dissonance cognitive, amenant à des ajustements et à un changement comportemental. Aussi, pour être efficace, le signal de « déviation » devrait notamment être saillant et fiable ; et l'efficacité du feedback serait supposée reliée à des caractéristiques propres au thérapeute, comme la perception de sa capacité à agir (la perception de contrôle - une auto-efficacité positive), son habileté à gérer des conséquences potentiellement négatives (que nous pouvons rapprocher des stratégies de coping), et une inclinaison et la motivation à atteindre le but ou des résultats (la valeur / ou dimension émotionnelle-motivationnelle). Dans ce modèle, si un individu s'écarte d'un objectif important à ses yeux, à condition qu'il perçoive le signal de feedback, il a la possibilité d'agir

de manière fonctionnelle (à condition de posséder une bonne auto-efficacité et de percevoir qu'il peut faire face aux conséquences de la situation). Le feedback n'est donc pas ici un simple phénomène répondant à un déterminisme mécaniciste.

Dans ce prolongement et revenant à des notions de cybernétique et d'ingénierie, Claiborn & Goodyear (2005) postulent que le feedback en pratique clinique s'appuie en général sur une intervention psychologique délibérée opérant deux fonctions basiques, qui sont retrouvées dans pratiquement tous les comportements humains liés à une interaction. Le feedback revêtirait ainsi une fonction d'information et une fonction d'influence ; et le feedback faciliterait le changement comportemental par l'influence interpersonnelle. Ces fonctions, au sein des interactions, se produisent en général naturellement et sans que les individus en aient la moindre conscience. La possibilité d'amener au niveau conscient ces éléments, ordinairement traités automatiquement, serait une des tâches de la psychothérapie. Le feedback du patient posséderait donc une fonction communicative, ou informationnelle (descriptive) d'un signal qui pourrait être verbal (ou non), et également une fonction évaluative ; ce signal pouvant être jaugé à la lueur d'un critère de performance (par exemple la cotation d'un questionnaire de mesure des résultats de la psychothérapie basé sur le feedback du patient). Il revêtirait également une fonction, potentiellement très importante selon nous, de dévoilement du contexte émotionnel de l'émetteur, et une fonction d'interprétation en regard de ce contexte ; le signal étant interprété en fonction d'un référentiel théorique commun, comme par exemple l'orientation psychothérapeutique dans laquelle s'inscrit le traitement, et qui va rendre plus ou moins importante la signification de ce signal. Ces dimensions pourraient être présentes de manière exclusive ou combinatoire. Les auteurs proposent également dans ce modèle que des variables liées aussi bien au thérapeute qu'au patient puissent influencer les effets dus au feedback (la crédibilité de la source, la valence du feedback avec une feedback négatif plus difficile à recevoir, l'humeur du receveur – une humeur plus basse étant associée à un feedback mieux reçu).

Dans la lignée des modèles théoriques antérieurs, la théorie du feedback contextualisé (The Contextual Feedback Intervention Theory) de Bickman et al. (2011), suggère que le feedback opérerait ainsi comme un signal qui aiderait à réduire l'écart entre un état désiré et une situation actuelle ; en précisant des co-variables, non prises en compte par les tentatives de modélisation précédentes. Pour augmenter les performances, le feedback devrait être recueilli ou présenté selon une fenêtre temporelle définie et apporter de nouvelles informations, notamment sur les erreurs commises (par exemple que le progrès n'est pas aussi bon que celui qui est attendu) ; et être associé à des stratégies de résolution de problème qui peuvent être effectivement déployées.

Pour synthétiser quelque peu ces modélisations, le feedback en psychothérapie pourrait apparaître comme un vecteur d'informations importantes, dont la révélation, en fonction des caractéristiques propres au thérapeute, permettrait d'influencer et donc de modifier ses propres comportements.

Dans cette optique, Sapyta, Riemer & Bickman (2005) opérationnalisent une définition des systèmes de mesure du feedback en psychothérapie (Measurement Feedback System - MFS), sur la base de deux composantes : une mesure ou une combinaison de mesures administrées durant le traitement, et agissant comme un système de collecte continu d'informations reliées au processus et aux résultats du traitement (nous retrouvons ici la dimension informative), et une composante de renvoi efficace (cliniquement et temporellement) de l'information aux cliniciens (la dimension d'influence).

Pour illustrer plus avant la diversité des tentatives de modélisation de ces mécanismes, et leurs points communs, nous pouvons également nous rapporter aux travaux des auteurs de la revue Cochrane sur l'efficacité de la pratique ROM (Kendrick et al., 2016). Citant Carlier (2012), ils avancent que les effets des interventions basées sur le feedback (FI) pourraient être expliquées par deux méta-théories : la théorie du FIT (Feedback Intervention Theory) et la théorie de l'évaluation thérapeutique (Therapeutic Assessment ou TA). Dans la théorie du FIT,

le feedback des résultats des patients obtenu par la passation d'échelles autoévaluées (PROMS) influencerait les cliniciens à ajuster leur traitement ou à guider les patients vers d'autres interventions alternatives ; ce qui améliorerait les soins. Ici encore la dimension de signal et d'influence sont donc prépondérantes. Dans la TA, ce serait la provision du feedback au patient, en plus de la provision au clinicien, qui amènerait à une cascade de conséquences positives en engageant les patients dans les soins, en facilitant une meilleure communication et une meilleure compréhension du contexte du patient, et en autorisant la prise de décision partagée entre les patients et les cliniciens. Ce qui aurait pour effet d'améliorer leur entente et l'adhérence du patient au traitement (orienté vers des buts communs). Ceci augmenterait la satisfaction du patient, ce qui serait finalement à même de potentialiser l'efficacité des soins. L'accent est donc également mis sur les mécanismes sociocognitifs amenant à un meilleur engagement des clients dans les soins. Effectivement, la mesure fréquente des symptômes par les patients serait de nature à augmenter leur prise de conscience et leur capacité à référer des éléments importants au clinicien ; ou leur donner le sentiment d'être mieux aidé (Dowrick, 2009). À ce sujet, Miller, Hubble & Duncan (2007) ont rapporté que les meilleurs résultats psychothérapeutiques (évalués par le système de gestion des résultats PCOMS) ont été retrouvés chez les thérapeutes capables d'obtenir du feedback négatif sur l'alliance thérapeutique dès le début du traitement.

1.3.2.2. La problématique de la rupture de l'alliance thérapeutique

Ainsi les éléments négatifs rapportés par les patients peuvent notamment véhiculer le signal d'une rupture de l'alliance, et dans la pratique ROM, le feedback peut donc être utilisé pour évaluer la qualité de la relation thérapeutique, créer une meilleure alliance ou réparer et prévenir les ruptures. Le travail doctoral de Laraway (2015) éclaire ce point en proposant un modèle rationnel-empirique des ruptures de l'alliance thérapeutique qui s'appuie sur l'analyse séance par séance du feedback du patient. Laraway s'appuie notamment sur le modèle de Safran et al. (2011) des ruptures de l'alliance, qui postule que la rupture de l'alliance peut être

considérée comme un comportement ou une communication des patients signalant des points critiques de la thérapie (et qui seraient à explorer). Ce phénomène très fréquent, et qui a été évalué dans certaines études comme se produisant dans 41 à 100% des séances quand elles sont analysées par des observateurs entraînés à les repérer, serait relié à de moins bons résultats positifs au traitement. La reconnaissance et la résolution de la rupture pourraient jouer un rôle significatif dans la rétention, l'engagement des clients dans les soins et donc, dans les résultats positifs à l'issue du traitement (Muran et al., 2009).

Précisant encore plus la nature des éléments négatifs qui pourraient être véhiculés par le feedback des patients, Laraway cite également les travaux de Hill, Thompson & Corbett (1992) qui ont démontré que les clients masquent ou cachent de manière consistante leurs réactions négatives dans les thérapies (brèves ou plus longues). Dans cette étude, 64% des patients n'ont pas rapporté volontairement certains éléments personnels pendant les séances, et la majorité de ces éléments ont revêtu un contenu négatif. Ceci nous renvoyant à une des limites des modèles de ruptures de l'alliance thérapeutique, basés sur la conscientisation, où les thérapeutes auraient la possibilité de reconnaître directement les phénomènes impliqués (notamment par une pratique réflexive). Il est pourtant clair que les thérapeutes présentent de grandes difficultés à percevoir le déclin de leurs patients au cours du traitement, et l'alliance thérapeutique évaluée par les patients est plus prédictive des résultats que le jugement du praticien (Horvath & Bedi, 2002). Également, les systèmes de ROM ont démontré une efficacité considérablement supérieure au jugement des cliniciens pour détecter les patients à risque de se détériorer. (Hannan et al., 2005).

1.3.2.3. Une synthèse des modèles explicatifs du feedback dans la pratique ROM

Des dimensions communes se dégagent donc de ces différentes modélisations des effets dus au feedback dans la pratique ROM : en premier lieu, la dimension communicationnelle du feedback qui agirait comme un signal, et qui apparaît comme un facteur commun de ces recherches ; suivie de la caractéristique d'influence du feedback recueilli qui est retrouvée selon

les modélisations, avec un effet aussi bien sur le thérapeute que sur le patient. Un consensus sur les résultats de l'implémentation de la pratique ROM semble également se dégager autour du fait que la provision du feedback instrumental (FI) n'aiderait que les cliniciens dont les patients ne retirent pas encore de bénéfices de leur traitement (les patients NOT). Ce feedback serait donc un signal d'influence efficace quand il véhiculerait une information négative. Il ne pourrait pas de ce fait constituer un signal de renforcement des thérapeutes, mais plutôt un signal « punitif » ou d'extinction, pour situer les processus impliqués dans une perspective comportementaliste. Il pourrait plutôt agir ainsi comme un correcteur ou un signal palliatif ; ce qui irait dans le sens du modèle en système GPS du feedback qui alerterait le navigateur en cas de déviation de la trajectoire idéale. Ce serait donc plus clairement un signal de feedback négatif qui exercerait une influence positive sur le traitement ou exercerait un effet tampon en diminuant les mauvais résultats attendus au traitement (par l'activation du thérapeute pour modifier la séquence de soins). Enfin, ce signal mettrait en lumière des éléments qui auraient été négligés aussi bien par le thérapeute que par le patient, et quelques caractéristiques reliées à l'efficacité du recueil de ce feedback en thérapie modéreraient aussi bien du côté patient que du côté thérapeute, l'efficacité de cette intervention : ainsi seraient retrouvés du côté du thérapeute, le genre (féminin avec un effet positif), l'auto-efficacité (positive), la propension au feedback (externe), la crédibilité de la source (autorité), l'humeur du récepteur (négative), la saillance (visibilité du feedback), la contiguïté temporelle et la fréquence (immédiateté et continuité), la forme (contextualisée avec un feedback accompagné d'une interprétation et d'un guide à l'action) en tant que variables importantes à prendre en compte dans les effets.

Enfin, il est clairement démontré que les thérapeutes n'apprennent pas du FI sans que des explications claires aient été apportées à ce phénomène. Nous pouvons de prime abord postuler que si les signaux de feedback artificiels (feedback instrumental) véhiculant le signal de « dysfonctionnement » du patient ou de la relation thérapeutique, ne sont pas présentés avec une fenêtre temporelle autorisant un apprentissage discriminatif, il n'est pas possible de

généraliser les acquis. En effet, en pratique clinique, les échelles sont souvent administrées au début et ou à la fin des séances, ne permettant pas de relier ou de « tracer » finement ce qui dysfonctionne en séance, avec le signal émis par le patient. Pour illustrer ce phénomène, nous allons revenir au modèle GPS de l'efficacité due au feedback. Dans une navigation guidée par un signal GPS, le conducteur a potentiellement l'occasion de saisir sur son trajet, ou pour utiliser des termes issus de l'analyse comportementale, de « discriminer », les signaux « publics », c'est-à-dire accessibles à tous, dans le monde extérieur, par l'usage des cinq sens. Par exemple, il peut prendre conscience d'un panneau signalant une déviation, remarquer du fait du craquement du bitume un trou béant sur la chaussée, ou déduire du bruit des klaxons qu'il se dirige vers un ralentissement notoire de la circulation. Si le moniteur GPS lui suggère un changement d'itinéraire, et que le navigateur ne peut pas, en même temps, percevoir le signal saillant et naturel (existant dans l'environnement) correspondant à la difficulté motivant un changement d'itinéraire (par exemple : une enfilade de véhicules signalant un bouchon), alors, il ne lui est pas possible d'apprendre que la voie est impraticable pour cette raison. D'autant plus qu'il manquerait d'expérience dans ces situations. Et à l'avenir, devant les mêmes conditions, il risquerait de persister dans son comportement, de ne pas changer d'itinéraire, et de s'engager dans une voie impraticable, s'il ne dispose plus du signal GPS correcteur. Dans les conditions cliniques de la séance de psychothérapie, la situation est quelque peu plus compliquée puisque une bonne partie des comportements sont des « actions » dites intrapersonnelles et donc « privées » (les pensées, les émotions, les sensations du patient) ; elles ne sont donc pas observables directement par le clinicien ; et ici aussi le manque de contact du clinicien avec les contingences réelles (la détresse du patient), qui pourrait être pourtant en partie réalisé par l'émission immédiate d'un signal de feedback (verbal) et d'indices fiables signalant la difficulté, ne permettent pas au clinicien d'apprendre de la situation pour généraliser.

Pour clarifier ces propos, partons par exemple, d'une perception négative du patient quant à l'aide du thérapeute : une mauvaise alliance thérapeutique, qui peut difficilement être exposée sans verbalisation ou sans un autre feedback (instrumental ou non) du patient. Si cette difficulté n'est pas verbalisée et que le clinicien reçoit à la fin de la séance un signal instrumental de feedback négatif relatif à l'entretien, il lui sera difficilement possible de relier cette information à des phénomènes qui se sont produits à des moments précis de la séance (les déterminants de ce feedback), notamment ses propres actions pendant la séance. Ces feedbacks pourraient pourtant lui permettre de reconnaître les indices présentés par le patient d'une mauvaise relation thérapeutique, au cours de cette séance, potentiellement exploitables dans des consultations futures. Aussi à moins de disposer de signaux visibles naturels saillants et fiables, comme par exemple l'émission d'une « véritable » expression faciale émotionnelle négative, pouvant être reliés à un feedback verbal ou instrumental immédiat du sentiment de ne pas être aidé, émis par le patient, il n'est donc pas possible pour le clinicien de transférer les fonctions aversives du signal instrumental négatif de feedback à des éléments observables en séance. En ce sens, il n'y a donc pas d'apprentissage discriminatif chez le clinicien et pas de généralisation possible de ces acquis à d'autres contextes ; comme de nouvelles séances où se présenteraient les mêmes difficultés identifiables par les mêmes signaux naturels (la même expression faciale).

Nous pouvons postuler que ce phénomène expliquerait pourquoi, si le Feedback Instrumental n'est pas présenté dans les conditions de contiguïté temporelle avec un signal naturel fiable et indicateur de mauvais résultat, potentiellement discriminable, il ne serait pas possible pour le thérapeute de réaliser l'apprentissage que ce signal naturel véhicule une information négative. Dans ce sens, les thérapeutes ne pourraient pas généraliser leurs acquis aux conditions dans lesquelles le signal de FI n'est plus présent. Et les thérapeutes n'apprendraient pas de leur exposition au feedback instrumental. Ce phénomène et sa compréhension pourraient amener à une nouvelle modélisation de l'influence des processus de feedback dans la pratique ROM et en psychothérapie

A ce sujet, il a été mis en évidence que l'efficacité des thérapeutes varie dans les recherches, avec certains thérapeutes qui sont naturellement plus efficaces sans disposer de feedback Instrumental (en disposant des seuls signaux naturels de feedback de l'interaction clinique). Dans l'étude contrôlée randomisée de thérapies de couple d'Anker, Duncan & Sparks (2009), quand le signal de FI était fourni aux collègues naturellement moins efficaces, ces derniers pouvaient atteindre de meilleurs résultats sans autre entraînement plus formel. Ceci suggérant que le feedback seul puisse être correcteur des performances ou venir en « remplacement » du feedback naturel pour lequel les thérapeutes auraient des problèmes de perception variables. Le feedback instrumental aurait ici un effet correcteur-additif ou palliatif du feedback naturel présent dans les conditions classiques de la psychothérapie.

Ainsi, au-delà de la simple hypothèse que l'efficacité des interventions liées au feedback serait modulée par des caractéristiques fixées (des constantes comme le genre, la propension au feedback) identifiées et associées aux thérapeutes et aux patients, nous émettons l'hypothèse que certains thérapeutes seraient très efficaces sans apport du FI, et ne bénéficieraient donc pas autant du FI que leurs collègues moins efficaces, parce qu'ils auraient un accès privilégié à une « dimension cachée » de l'information disponible pendant les séances de psychothérapie. C'est à dire accès à une meilleure perception de ces éléments masqués ou passant inaperçus, en cours du traitement, et dont la prise en compte serait de nature à ré-orienter le processus des soins.

1.3.3. De l'existence de signaux d'alerte fiables de détérioration des patients et d'une mauvaise relation thérapeutique en situation naturelle de psychothérapie

Pour tenter de répondre à cette question, il est nécessaire de dépasser le cadre restrictif de la pratique clinique et des modélisations initiales du feedback, pour restituer le phénomène de la psychothérapie, dans une perspective transdisciplinaire ; en croisant des données issues des sciences de l'information, des sciences affectives, et des sciences évolutionnistes.

1.3.3.1. Une forme spécifique d'interaction et d'échanges verbaux et non verbaux

En effet comme le soulignent Orlinski et al. (2004) citant Frank et Frank (1993), la psychothérapie est essentiellement une forme spécifique de relation d'influence entre un thérapeute et son patient, où les capacités du soignant sont censées susciter les forces de guérison chez le patient. Dans cet échange, les patients et les thérapeutes partagent des signaux verbaux ; mais aussi des signaux non verbaux, pour laquelle Hall a proposé l'expression « dimension cachée » (1969). Pour Knapp (1972), il s'agirait de la communication qui ne comprend pas les mots écrits ou parlés. Pour Matsumoto et al. (2013), la communication non verbale (CNVB) représenterait le transfert et l'échanges de messages dans toutes les modalités qui excluent les mots. En psychothérapie pourtant, l'accent est principalement porté sur les mots et la pratique est communément désignée par le terme « talking cure » ou traitement basé sur la parole ; en raison d'un héritage freudien conséquent où le but du traitement a été historiquement d'amener à conscientiser des éléments « enfouis » chez le patient par le biais du langage. De ce fait, des efforts considérables de recherche ont été portés au cours du siècle dernier sur l'analyse du contenu verbal des entretiens afin de mieux saisir les déterminants de l'efficacité, et ceci dans tous les courants psychothérapeutiques ; au détriment de l'analyse et de l'étude du contenu para-verbal ou non verbal. Alors que la communication non verbale est à même de véhiculer des informations très importantes sur le processus de la thérapie ou sur l'état du thérapeute et de son patient. De manière inquiétante cet agenda de recherche centré sur le contenu verbal desservirait la recherche en opérant un réductionnisme dysfonctionnel des déterminants étudiés. En particulier, pour Friedman (1978), la vaste majorité des messages communiqués dans une interaction seraient de nature non verbale ; suivant les études de 65% à 95% des informations ne seraient pas simplement véhiculées par les mots. En particulier, la posture et les expressions faciales, qui apparaissent spontanément au cours d'un entretien, ont également une valeur communicative reliée au comportement verbal, observable chez des observateurs non entraînés. Ainsi alors que les thérapeutes peuvent facilement s'appuyer sur les signaux verbaux, ils

peuvent aussi retirer de l'information des signaux non verbaux (Ekman, O'Sullivan & Frank, 1999) et ces derniers peuvent être étudiés.

1.3.3.2. La fonction des messages non verbaux

Ces messages non-verbaux peuvent être exprimés par tous le corps, et même véhiculés par le contexte de l'interaction au sens large du terme (par exemple : l'habillement, le port de bijoux) ; un media particulier de communication se dégagerait cependant : le visage et ses expressions. En effet, pour Matsumoto & Hwang. (2013), le visage renverrait des signaux automatiques (et donc non ou peu filtrés) avec une temporalité immédiate, notamment des expressions émotionnelles. Les émotions étant considérées dans une perspective évolutionniste comme des mécanismes adaptatifs qui servent de nombreuses fins ; ce sont des réactions biopsychosociales transitoires aux évènements qui peuvent entraîner des conséquences pour notre bien-être, et qui permettent notamment de traiter rapidement de l'information, avec le minimum d'efforts conscients pour déclencher des réactions immédiates, en nous préparant à l'action. De plus, il est communément admis que les émotions soient universelles et qu'elles renvoient à des expressions faciales identifiables qui seraient également universelles ; elles fourniraient des informations fiables quant aux individus qui les émettent (notamment en signalant leur état affectif et leur intention), et joueraient donc un rôle critique sculpté par l'évolution, dans les interactions sociales (Matsumoto & Hwang., 2013). Et en particulier dans le contexte d'interaction face à face que peut représenter la psychothérapie. Elles évoqueraient aussi bien des comportements de réponse émotionnelle complémentaire chez autrui, comme une réponse d'apaisement du thérapeute face à un sentiment de détresse présenté par le patient, qu'elles véhiculeraient des informations sur l'état mental et les intentions de l'émetteur ; en particulier dans les interactions dyadiques (Keltner & Haidt, 2001). La capacité empathique, c'est-à-dire la capacité à s'imaginer « à la place de l'autre » est notamment explicable par les mécanismes automatiques d'imitation des expressions faciales émotionnelles (*mimicries* et neurones miroirs) et les thérapeutes qui sont les plus à même de détecter les émotions sont jugés

plus empathiques par leurs patients. Cette capacité est notamment altérée si les mouvements faciaux sont bloqués (Jospe, Flöel & Lavidor, 2018).

Alors que les émotions négatives auraient plutôt tendance à activer la branche sympathique du système nerveux central, c'est-à-dire à inhiber et restreindre le répertoire comportemental, les émotions positives seraient associées à un élargissement du champ d'action et à de meilleures connexions sociales (pour une revue voir Niedenthal et al., 2018, ch4.). Dans la relation thérapeutique, pour le patient, expérimenter des émotions négatives pourrait donc être associé chez le thérapeute empathique à des réactions de compensation, mais aussi à une restriction innée du répertoire comportemental et à une réponse plus rigide. En outre même si l'expression des émotions est probablement le plus important stimulus de la vie sociale, véhiculant une quantité d'information considérable sur leur émetteur, par le biais du visage comme canal privilégié, le reste du corps et la voix sont également des canaux très importants qui permettent notamment d'intégrer les informations dans leur contexte (Niedenthal et al., 2018).

En psychothérapie, les émotions ont donc une importance considérable, en étant considérée comme des processus thérapeutiques (les prototypiques expériences émotionnelles correctrices de Alexander & French (1946) liées aux réactions de compensation des thérapeutes), mais aussi parce qu'elles apportent des informations fiables sur les processus intra-individuels présents chez les patients et sur la qualité de la relation thérapeutique, non disponibles par ailleurs. L'empathie a d'ailleurs été reliée à des meilleurs résultats en psychothérapie (Wampold & Imel, 2015) et la capacité à se mettre à la place de l'autre est rendue possible par les capacités de prise de perspective (ToM, pour une revue du développement voir Wellman, 2018). Mais aussi par la capacité à reproduire les expressions faciales pour mieux ressentir ce qu'expérimente l'autre (théorie initialement développée par Théodore Lips (1851-1914)).

1.3.3.3. La capacité à décoder les messages non verbaux.

Si l'expression des émotions peut paraître universelle, leur modulation est influencée par les contextes culturels et sociaux, ce qui peut interroger sur la fiabilité des expressions émotionnelles dans le contexte de la psychothérapie où les patients tendent à dissimuler des informations négatives et importantes. La reconnaissance des émotions est également considérée comme universelle (Matsumoto & Hwang., 2013) et les psychothérapeutes sont donc naturellement à même de discerner les expressions émotionnelles présentes chez leur patient, qui renvoient aux sept émotions primaires mises en évidence par Ekman (2003) souvent appelées macro-émotions. Ils sont capables d'accéder à une partie de leurs ressentis. Il existerait pourtant, des différences notables dans la capacité des individus à distinguer les plus petites expressions émotionnelles dénommées micro-émotions ou émotions subtiles. En effet, il a été démontré que les émotions qui apparaissent sur le visage dans des contextes sociaux où il n'est pas nécessaire de modifier ou de censurer l'expression ont tendance à durer plus longtemps et à s'exprimer sur l'entièreté du visage (pendant 0,5 à 4 s), tandis que quand le contexte social nécessite de dissimuler les émotions, leur expression s'afficherait avec une moindre intensité, ou sur une partie seulement du visage, et pendant des durées beaucoup plus faibles (moins de 0,5 s ; Matsumoto & Hwang, 2013, Ekman, 2003). Les macro-expressions seraient donc beaucoup plus faciles à discriminer que les micro-expressions qui sont très rapides et difficilement identifiables pour la plupart des individus. Pourtant ces petites expressions véhiculeraient des signaux plus fiables de l'état émotionnel d'un individu que les macro-expressions qui, peuvent être manipulées, quand un individu chercherait à dissimuler son état ou quand il n'en aurait pas conscience. Ainsi, de manière générale, pour Ekman & Friesen (1969), la communication non verbale pourrait venir se substituer complètement à la communication verbale (et agirait selon nous comme un palliatif), la compléter (agissant comme un correcteur) ou la contredire (réalisant une correction), et ainsi jouer un rôle déterminant en fournissant des éléments à prendre en compte par les thérapeutes pour décider

de leurs actions. En ce sens, le feedback instrumental renvoyé par les systèmes ROM pourrait même être considéré comme une extension de la communication non-verbale, quand ces systèmes permettent de véhiculer des messages tels que des signaux de détérioration ou un graphique de progression à même d'introduire une réorientation de la thérapie.

1.3.3.4. Les signaux non verbaux fiables : micro expressions et expressions subtiles

Dans le prolongement des travaux de Duchenne et Darwin sur l'anatomie de l'expression émotionnelle, Haggard & Isaacs (1966) ont initialement mis en évidence l'existence et la présence de micro-expressions émotionnelles difficilement discernables pour l'œil non entraîné, en analysant des enregistrements vidéo de psychothérapie image par image. Des micro-expressions furent ensuite retrouvées chez des patients dépressifs filmés et analysés par Ekman et Friesen (1974). Plus récemment, ces travaux ont été popularisés par Paul Ekman et portés à l'écran dans la série télévisée américaine « Lie To Me » inspirée de sa vie et de son travail, où le personnage principal est chargé de détecter le mensonge dans des contextes d'interventions très divers. De manière plus spécifique pour notre propos, en 2008, Porter et Ten Brinke publièrent la première recherche démontrant la présence de micro-expressions émotionnelles chez des individus qui essayaient de masquer leurs émotions. Puisque ces micro-expressions sont potentiellement le signe d'émotions dissimulées, et que les recherches ont montré que les patients avaient tendance à cacher du contenu négatif pendant les séances de psychothérapie, ces signaux pourraient véhiculer des feedbacks naturalistes potentiellement très importants pour les thérapeutes. Et la différence dans l'habilité à détecter ces informations pourrait expliquer une part de la variance des résultats des meilleurs thérapeutes et une part de la variance des résultats dans les interventions utilisant le feedback instrumental en psychothérapie. Les praticiens qui ont la possibilité de se saisir de ces informations, étant mieux armés pour adapter le processus thérapeutique. En particulier, les recherches ont démontré qu'il existait un lien entre la capacité à reconnaître les émotions subtiles (qui sont souvent des manifestations d'expressions faciales émotionnelles partielles) et la capacité à détecter la

tromperie (Warren, Schertler & Bull, 2009). La plupart des individus ne sont pas efficaces dans la détection des émotions micro et subtiles, et certains individus sont naturellement plus talentueux. Ekman, O'Sullivan & Frank (1999) ont publié la première étude démontrant que certains groupes de professionnels, plus spécialement des psychologues intéressés par le phénomène de la détection du mensonge, pouvaient obtenir de meilleures performances à détecter les menteurs sur des enregistrements vidéo. Ce qui les distinguerait de leurs congénères qui, eux, sont plus généralement capables de reconnaître aisément uniquement les macro-expressions faciales. Que ce soit à la suite d'un entraînement spécifique ou du fait d'une capacité innée, les femmes ont été identifiées comme souvent plus efficaces dans l'interprétation des signaux non verbaux (Hall, 1978), et comme nous l'avons vu précédemment, elles sont également plus efficaces dans l'utilisation du feedback instrumental.

De plus, nous l'avons évoqué, les individus tendent à moduler l'expression de leurs émotions dans les contextes sociaux (Matsumoto & Willingham, 2006) et en thérapie (Hill, Thompson & Corbett, 1992). En séances de psychothérapie, les macro-expressions pourraient ainsi être modulées alors que les micros-expressions ou les expressions subtiles ne pourraient pas l'être ou difficilement ; ce qui donnerait un avantage certains aux thérapeutes qui, en condition naturaliste, sont plus à même de détecter ou de décoder les signaux non verbaux. L'entraînement à ces habiletés est d'ailleurs documenté et permettrait une utilisation dans un environnement de production (Hwang & Matsumoto, 2010).

Ainsi, dans un travail de mémoire de master, Sutter (2010), a émis l'hypothèse que l'entraînement à la reconnaissance des signaux non verbaux (les micro-expressions faciales émotionnelles) serait de nature à augmenter la capacité des thérapeutes à identifier les émotions, et renforcerait de ce fait l'alliance thérapeutique. Nous pouvons rappeler à ce sujet qu'une des trois dimensions de l'alliance thérapeutique telle qu'elle a été définie trans-théoriquement par Bordin (1979) est la capacité du thérapeute à comprendre son patient, qualifiant la qualité de la relation. Les résultats expérimentaux ont laissé effectivement apparaître que les thérapeutes qui

avaient suivi une formation à la reconnaissance de ces micro-expressions, obtenaient ensuite de meilleurs scores d'alliance thérapeutique évalués par l'échelle SRS (Session Rating Scale, Miller et al., 2005) ; suggérant un lien entre la reconnaissance du comportement non verbal, en particulier des micro-expressions, et une meilleure alliance entre le thérapeute et son patient. De meilleurs résultats au traitement mesurés par l'échelle ORS (Outcome Rating Scale ; Miller et al., 2005) ont été également retrouvés dans cette recherche, suggérant un lien direct entre une meilleure perception des micro-émotions et de meilleurs résultats psychothérapeutiques.

1.3.4. Vers un modèle explicatif des effets dus au feedback dans la pratique ROM

Toutes ces données soulignent le potentiel de dévoilement des informations apportées par les signaux non verbaux émis par les patients (ou feedback naturalistes). Les thérapeutes sont naturellement habilités, à moins de présenter des déficits cognitifs ou neurologiques marqués, à décoder les signaux verbaux émis par leurs patients, au moins dans leur langue maternelle, puisque le langage articulé est la base de la communication orale. Mais aussi à décoder des signaux non-verbaux, en particulier à reconnaître des expressions faciales émotionnelles avec une efficacité variable. Ces éléments, en lien avec les modélisations classiques des effets dus aux interventions basées sur le feedback, nous permettent de formuler quelques hypothèses.

1.3.4.1. Du feedback non verbal au feedback instrumental

Comme les capacités à distinguer les signaux non verbaux véhiculant des informations sur l'état des patients diffèrent selon les thérapeutes, les thérapeutes les mieux armés à distinguer ces signaux pourraient potentiellement obtenir de meilleurs résultats pendant les soins (les résultats préliminaires de la littérature semblent aller dans ce sens). Ensuite, les meilleurs thérapeutes distingueraient plus facilement les signaux qui sont les plus difficilement perceptibles pour la plupart des êtres humains, mais pour lesquels la fiabilité est mieux documentée (les micro-expressions et les expressions subtiles). La capacité à détecter les

émotions négatives masquées par les patients pourrait expliquer pourquoi le feedback instrumental négatif amènerait à de meilleurs résultats chez des patients qui initialement ne s'amélioraient pas et pour lesquels les thérapeutes auraient de grandes difficultés d'identification. Ces émotions négatives entraînant par ailleurs des réactions de compensation chez les thérapeutes auto-efficaces qui sont susceptibles d'amener à des changements dans la prise en charge et d'améliorer l'état du patient ; mais également des réactions défensives (comme ignorer le signal de feedback) pouvant bloquer la réaction fonctionnelle due à la perception d'un signal de détresse par les thérapeutes manquant d'auto-efficacité. Par ailleurs, la meilleure efficacité du FI est notamment liée à une humeur des thérapeutes plutôt congruente ; et l'utilisation du feedback est potentiellement liée à une réaction de surprise chez les thérapeutes (Chow, 2014). Cet effet de surprise serait à même, d'amener à de nouvelles réactions chez le thérapeute puisqu'il viendrait contrecarrer le pouvoir d'offuscation au feedback que représente la force de l'habitude qui rend insensible au contexte.

Puisque le feedback instrumental agirait pour aider les thérapeutes qui n'ont pas détecté les signaux de détérioration des patients dans les conditions NOT, les praticiens qui seraient à même de distinguer plus facilement les signaux négatifs et de surcroît les micro-signaux émotionnels négatifs seraient les plus efficaces. Puisque certains groupes de praticiens sont naturellement plus à même de lire ces signaux, nous pouvons émettre l'hypothèse que les meilleurs thérapeutes puissent être les plus efficaces à détecter ces signaux. Le feedback instrumental remplacerait alors (hypothèse palliative) ou compléterait (hypothèse supplétive) un processus discriminatif naturel qui serait à l'œuvre dans la thérapie, et pour lequel les praticiens varieraient en capacité. Les plus performants auraient accès à ces informations naturelles qui leur permettraient de mieux répondre aux enjeux de la relation thérapeutique en temps réel.

D'un point de vue expérimental, de mauvais résultats aux traitements se manifesteraient par l'émission chez le patient de signaux non verbaux (naturels) et auto-rapportés (feedback

instrumental) négatifs signalant la détresse émotionnelle (tristesse, peur) ou d'autres émotions négatives reliées au mauvais processus thérapeutique (colère, dégoût, mépris).

Nous pourrions supposer que l'effet direct d'une communication non verbale négative sur les résultats serait à priori négatif. L'alliance prédisant les résultats dans la littérature, l'effet de la communication non verbale sur les résultats pourrait transiter par l'alliance thérapeutique qui réaliserait ainsi une médiation de la relation entre la communication non verbale et les résultats.

1.3.4.2. Une tentative de modélisation du feedback instrumental dans la pratique ROM.

Comme le Feedback instrumental semble influencer les résultats seulement en condition NOT, nous pouvons supposer que le feedback instrumental puisse agir en médiant les effets négatifs de la communication non-verbale (micro et subtile notamment) en condition NOT sur l'alliance thérapeutique ou en modulant (hypothèse de modération) les effets négatifs de la communication non-verbale (micro et subtile notamment) en condition NOT sur l'alliance thérapeutique.

Cette modélisation est évidemment simpliste et réductrice tant les variables en interaction sont nombreuses ; de plus, les interactions du thérapeute et du patient avec les signaux de feedback ne sont évidemment pas synchroniques (analysables en transversal) mais sont diachroniques/séquentielles (analysables en longitudinal), du fait des boucles de rétroactions. Néanmoins, cette théorisation permet de dessiner un modèle structural illustrant des relations entre les construits latents (Alliance thérapeutique, Résultats, Communication Non verbale, et Feedback Instrumental) qui peuvent être opérationnalisées avec des indicateurs obtenus par une mesure (mesure de l'Alliance, du feedback, des émotions, des capacités de décodage des messages non verbaux des thérapeutes, mesure des résultats de la thérapie). Et ce modèle pourrait être potentiellement soumis à l'épreuve des faits à l'aide de traitements statistiques.

Le FI pourrait ainsi faire transiter des signaux négatifs non-verbaux réalisant une médiation entre les signaux non verbaux (macro, micro et subtils), l'alliance thérapeutique et les résultats (Figure 1). Nous retrouverions ici l'effet « GPS » où le feedback renverrait un signal négatif fiable qui pallierait les capacités naturelles de navigation du thérapeute (de détection des signaux non verbaux), lui permettant de réorienter la thérapie, et impactant l'alliance. Dans le modèle palliatif du signal de feedback Instrumental, le FI pourrait donc éventuellement agir en faisant transiter tout l'effet de la communication non verbale négative et représenter un remplaçant des signaux de communication non verbale réalisant une médiation totale : ceci correspondrait à la situation de navigation GPS typique où le navigateur s'appuie totalement sur le feedback de l'instrument de navigation pour poursuivre son chemin.

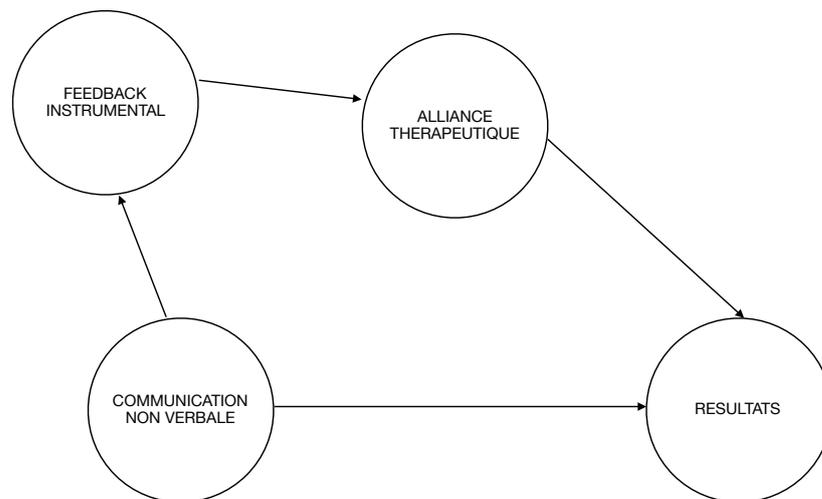


Figure 1 : Modèle palliatif du feedback - Modèle en médiation

Le FI pourrait aussi modérer la médiation entre les signaux négatifs non verbaux (macro, micro et subtils), l'alliance thérapeutique et les résultats (Figure 2). Nous pourrions ainsi imaginer l'effet « clinique » du feedback négatif en psychothérapie par un effet de lentille (Lens Effect), où le fait de recevoir du feedback négatif instrumental potentialiserait (focaliserait) la réaction des cliniciens, alors que des signaux non verbaux négatifs sont émis, et la réaction du thérapeute au feedback, corrigerait l'effet des signaux non verbaux négatifs impactant de ce fait positivement l'alliance thérapeutique. Dans ce modèle supplétif/correctif du feedback Instrumental (Lens model), le FI modulerait l'effet des signaux de communication non verbale.

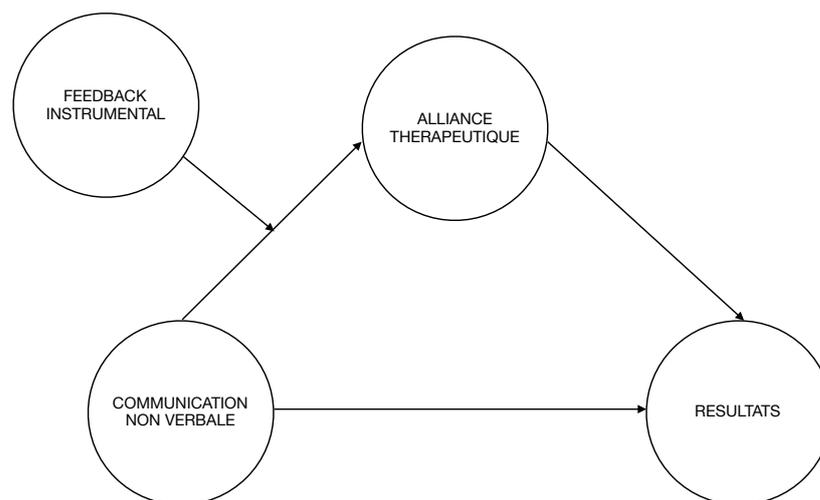


Figure 2 : Modèle supplétif du feedback - Modèle en modération

Au niveau expérimental, en opérationnalisant les construits du modèle palliatif, nous pourrions supposer qu'en condition NOT (de mauvais résultats), le feedback instrumental négatif exercerait un effet négatif correcteur sur l'alliance thérapeutique (de mauvaise qualité), l'alliance thérapeutique étant classiquement dans la littérature reliée positivement aux résultats

du traitement (Figure 3). Ici le rôle de la communication non verbale négative sur les mauvais résultats est en ce sens direct.

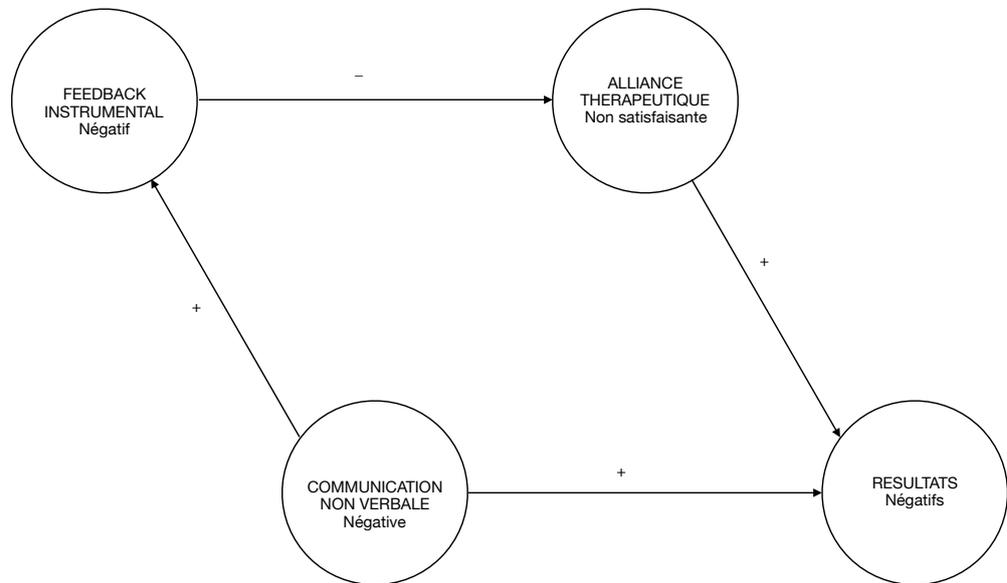


Figure 3. Hypothèse de l'effet médiateur palliatif du feedback négatif sur l'alliance thérapeutique dans la condition NOT (mauvais résultats)

Au niveau expérimental, en opérationnalisant les construits du modèle supplétif, nous pourrions supposer qu'en condition NOT, le feedback instrumental négatif exercerait un effet négatif et donc d'affaiblissement, ou de diffraction du signal pour reprendre la métaphore du modèle de la lentille (un effet de tampon), sur la relation positive entre les signaux non verbaux (négatifs) et la mauvaise alliance thérapeutique, l'alliance thérapeutique étant classiquement dans la littérature reliée positivement aux résultats du traitement, qui sont ici négatifs (Figure 4). Ici le rôle de la communication non verbale négatif sur les mauvais résultats est également en sens direct.

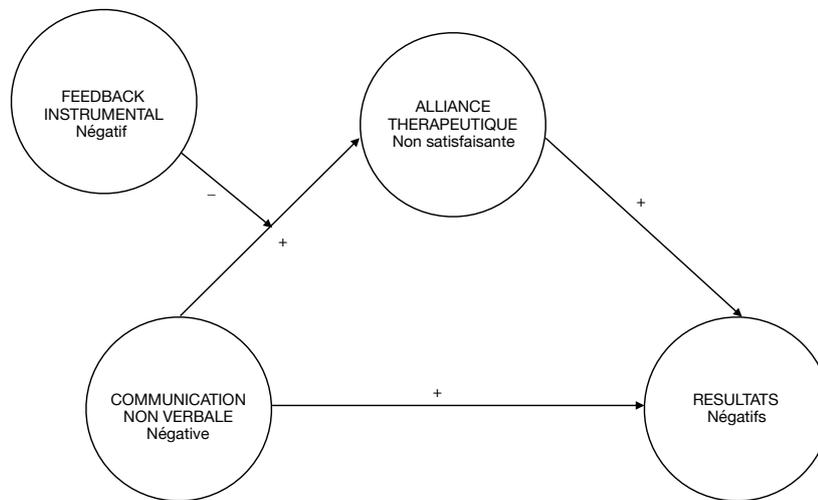


Figure 4. Hypothèse de l'effet modérateur / supplétif du feedback négatif sur l'alliance thérapeutique dans la condition NOT (mauvais résultats)

1.3.4.3. Conjectures liées au modèle palliatif du FI (modèle de médiation)

Nous pouvons formuler les suppositions suivantes qui sont articulées au modèle palliatif des effets dus au feedback instrumental dans la pratique ROM :

1) Dans les conditions naturelles de la psychothérapie (hors feedback instrumental), des différences significatives de capacité de reconnaissance des signaux non verbaux existent entre les thérapeutes les plus efficaces et les moins efficaces. 2) En conditions naturelles (hors feedback instrumental), des différences significatives d'efficacité existent entre les thérapeutes entraînés à repérer les signaux de CNVB et non entraînés à repérer les signaux de CNVB. 3) Les thérapeutes les plus efficaces dans la détection des signaux non verbaux négatifs sont aussi les plus efficaces à prédire les mauvais résultats. 4) Il n'existe pas de différences significatives d'efficacité entre les thérapeutes qui sont entraînés à repérer les indices non verbaux (entraînement CNVB) et les thérapeutes entraînés et ayant accès au feedback instrumental

(entraînement FI). Le FI agirait donc comme un remplaçant du signal non verbal (Hypothèse palliative)¹⁰. 5) Il n'existe pas de différence d'efficacité entre les thérapeutes en condition FI plus entraînement à la détection de la communication non verbale et les thérapeutes en condition FI seule (l'entraînement à la détection de la communication non verbale n'améliore pas les performances en condition FI). 6) En conditions naturelles (hors Feedback Instrumental), les thérapeutes les plus efficaces dans la perception des micro-expressions faciales (en particulier négatives) seraient les plus efficaces.

1.3.4.4. Conjectures reliées au modèle supplétif du FI (modèle de modération) :

1') Dans les conditions feedback instrumental, des différences significatives de reconnaissance des signaux non verbaux existent entre les thérapeutes les plus efficaces et les moins efficaces. 2') Dans les conditions feedback instrumental, des différences significatives d'efficacité existent entre les thérapeutes entraînés à repérer les signaux de CNVB et les thérapeutes non entraînés à repérer les signaux CNVB. 3') En condition de feedback instrumental, les thérapeutes les plus efficaces dans la détection des signaux non verbaux négatifs sont aussi les plus efficaces à prédire les mauvais résultats. En condition de feedback instrumental, les thérapeutes les plus efficaces dans la détection des signaux non verbaux négatifs sont aussi les plus efficaces. 4') Il existe une différence significative d'efficacité entre les thérapeutes qui sont entraînés à repérer les indices non verbaux (entraînement CNVB) et les thérapeutes entraînés et ayant accès au feedback instrumental (entraînement FI). Le FI agirait donc comme un complément au signal non verbal (Hypothèse supplétive). 5') Il existe une différence d'efficacité entre les thérapeutes en condition FI plus entraînement à la détection de

¹⁰ Ces hypothèses ont été présentées au Dr Paul Ekman (communication personnelle, 4 novembre 2013), dans le cadre de l'élaboration d'un premier projet de recherche sur la modélisation des effets du feedback en psychothérapie, projet qui n'a malheureusement pas pu être mené à son terme comme nous l'évoquons plus avant à la section 2.2 de ce travail.

la communication non verbale et les thérapeutes en condition FI seule. 6') En conditions de Feedback Instrumental, les thérapeutes les plus efficaces pour distinguer les micro-expressions faciales (en particulier négatives) seraient les plus efficaces.

Nous avons passé en revue dans cette partie, les différentes tentatives de modélisation de l'effet du feedback dans la pratique ROM en psychothérapie. Il apparaît notamment que le feedback instrumental agirait plus particulièrement dans les conditions de soins où les patients ne s'amélioreraient pas initialement (patients NOT), en signalant les mauvais résultats aux psychothérapeutes et aux patients, afin de les aider à adapter le traitement. D'autres signaux, non verbaux, de la mauvaise réponse au traitement, ou de la mauvaise relation thérapeutique, disponibles en condition naturaliste de psychothérapie, pourraient être liés à l'efficacité du feedback instrumental en condition NOT, et à l'efficacité de la psychothérapie en général. En pratique clinique, le feedback instrumental pourrait avoir un effet de lentille grossissante (hypothèse supplétive), permettant aux cliniciens de mieux percevoir les informations négatives sur les résultats des soins qu'ils avaient négligés, afin de mieux adapter le traitement. De manière expérimentale, la provision du feedback instrumental au clinicien modulerait en l'affaiblissant, l'impact négatif d'une mauvaise communication (identifiée par des signaux non verbaux) sur l'alliance thérapeutique entre le patient et le thérapeute qui, elle, agirait comme un médiateur des résultats. Le feedback agirait plutôt comme un complément d'information (effet loupe ou lentille-modèle lens) que comme un remplaçant des signaux présents naturellement (Modèle du GPS).

Plusieurs conjectures reliées à l'opérationnalisation de ce modèle sont proposées, dans le but de stimuler de nouvelles recherches sur la capacité des thérapeutes à discriminer les signaux non verbaux émis par leurs patients en condition feedback et non feedback, en particulier les signaux négatifs ; à étudier l'impact associé en terme de résultats psychothérapeutiques et également à comparer les effets d'un entraînement à la reconnaissance des signaux non verbaux en comparaison d'un entraînement aux pratiques de feedback

instrumental de la ROM. Ces recherches seraient à même, selon nous, d'éclairer les effets du feedback instrumental en psychothérapie, de mieux élucider les effets dus au traitement en général et de proposer de nouvelles pistes dans l'entraînement des psychothérapeutes susceptibles d'améliorer significativement leurs résultats.

Dans la section finale de cette première partie théorique liée à l'évaluation et à l'amélioration de l'efficacité des psychothérapies, aux apports du Routine Outcome Monitoring, à l'entraînement dit de pratique délibérée du thérapeute, et aux tentatives de modélisation des effets dus au feedback instrumental, nous allons rassembler les variables pertinentes qui ont pu être identifiées. Nous concluons cette section en formulant des hypothèses de recherche qui seront soumises à l'épreuve des faits dans la section empirique de ce travail de thèse.

1.3.5. Synthèse des modèles relatifs à l'utilisation de la pratique ROM en psychothérapie

:

- Le feedback ROM agirait comme un signal correcteur négatif qui influencerait le thérapeute (et éventuellement le patient) dans la course de la thérapie.
- Des modérateurs associés aux caractéristiques des thérapeutes, comme leur propension à recevoir du feedback et leur capacité à gérer les informations négatives (auto-efficacité) moduleraient les gains dus au feedback.
- Les thérapeutes ne pourraient pas apprendre de leur exposition aux systèmes de gestion des résultats basés sur le feedback des patients, car ils ne pourraient pas associer dans une fenêtre temporelle efficace, ces feedbacks instrumentaux avec les informations négatives naturellement disponibles dans les interactions thérapeutiques.
- Les thérapeutes pourraient augmenter sensiblement leur efficacité par la provision d'un feedback instrumental négatif sans autre entraînement formel ; les thérapeutes initialement les moins efficaces étant les plus aidés par le feedback.
- Le feedback instrumental pourrait agir en remplacement ou en complément de feedback naturels non verbaux présents dans les consultations, et pour lesquels les thérapeutes auraient des difficultés de perceptions variables.

Deux modèles des effets dus au feedback dans la pratique ROM sont proposés, ainsi que plusieurs conjectures susceptibles d'être éprouvées dans de nouvelles recherches.

1.4. Synthèse de la partie théorique et problématique de la recherche

1.4.1. Synthèse des variables d'intérêt des effets dus au feedback instrumental dans la pratique ROM.

Dans cette première partie de notre travail, nous avons parcouru les aspects théoriques qui selon nous pourraient éclairer le fonctionnement et l'efficacité du feedback instrumental en psychothérapie. Nous allons passer en revue ces variables d'intérêt pour les reprendre et les articuler avec un schéma récapitulatif.

Ainsi, nous avons pu évoquer que la psychothérapie était avant tout une méthode de changement relationnelle qui fonctionnait comme une communication d'influence entre un thérapeute et son patient. Dans cette interaction, les thérapeutes pouvaient aussi bien échanger des informations de nature verbale que des informations de nature non verbale. Cette communication thérapeutique ayant pour résultante la création d'une bonne alliance thérapeutique vers des buts communs, et l'engagement des clients dans les soins, c'est-à-dire dans les tâches du traitement, afin de produire des résultats. Ce processus n'opérant pas de manière mécaniciste, mais à l'aide de boucles de rétroaction, entre le thérapeute et son patient. Les signaux de rétroaction pouvant être constitués de comportements verbaux ou non verbaux, de feedback naturaliste ou de feedback recueillis au travers d'un instrument de mesure.

Le patient, le thérapeute et leurs caractéristiques respectives sont également en interaction et peuvent modérer, dans la production des résultats, les réactions dues aux processus de feedback. Du côté du patient, le degré de détresse initiale à l'entrée dans les soins est un prédicteur important des résultats au traitement. Du côté du thérapeute, le genre, la propension à recueillir du feedback, l'humeur et l'auto-efficacité, modulent ces interactions et les réactions du thérapeute.

Enfin, dans la situation d'une thérapie orientée par le feedback (FIT), nous avons postulé que le feedback instrumental pourrait influencer et interagir avec les feedbacks naturalistes

présents dans les communications thérapeutiques. L'effet de ce feedback pouvant être lui-même modulé par les caractéristiques intrinsèques du feedback instrumental (saillance, fréquence, contenu).

La figure 5 reprend donc et articule ces différents déterminants qui ont été mis en évidence dans l'analyse de la littérature des effets dus au feedback dans la pratique ROM, avec l'addition de l'hypothèse supplétive ou de médiation du feedback instrumental.

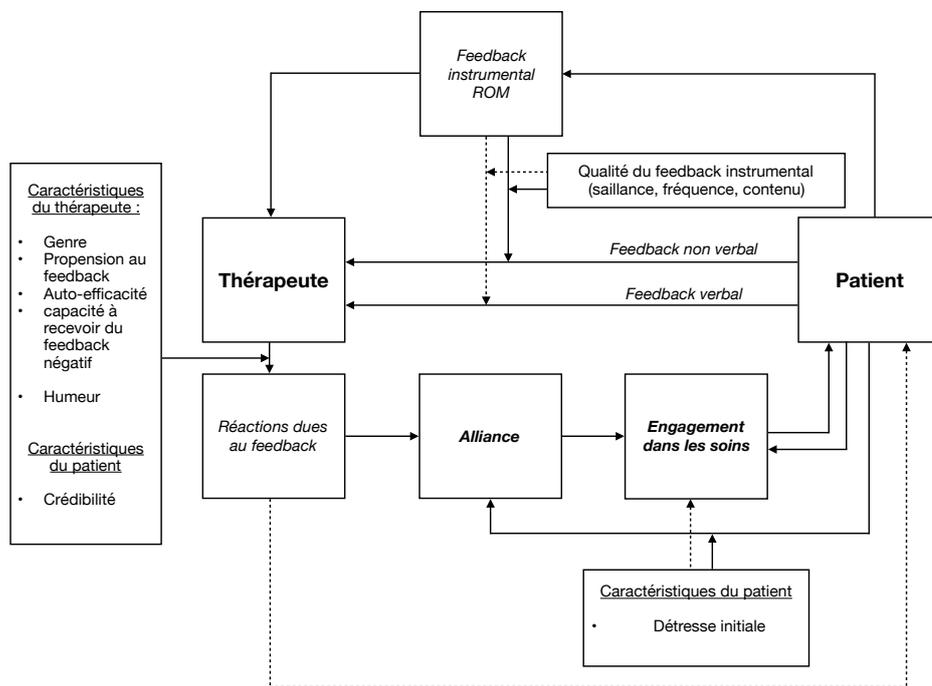


Figure 5. Tentative de modélisation du feedback dans la pratique ROM

1.4.2. Hypothèses relatives à l'adaptation française d'un système de Routine Outcome Monitoring : le système PCOMS

Dans le but de stimuler les recherches et la pratique dans l'utilisation du Routine Outcome Monitoring (ROM) dans les contextes francophones, nous avons décidé de réaliser l'adaptation transculturelle des échelles ORS et SRS du système PCOMS (Miller et al, 2005) enregistré au répertoire américain de la SAMSHA. Une telle validation nous semble préalable à son utilisation dans les recherches recourant à un système ROM.

L'objectif de cette première étude est donc d'élaborer une version française des échelles ORS et SRS du système PCOMS (Miller et al, 2005) et d'en vérifier les qualités psychométriques afin de permettre à la communauté de psychothérapeutes francophones de disposer d'un système ROM facile d'utilisation.

A cet effet, nous examinerons a) la fiabilité (c'est-à-dire la cohérence interne des scores obtenus aux items des échelles ORS et SRS, b) la structure factorielle des échelles au moyen d'une analyse factorielle de confirmation (AFC), et c) la validité de construit de ces questionnaires.

Nous posons donc l'hypothèse générique suivante quant à la validation française de ces échelles : une hypothèse nulle d'équivalence en termes de qualités psychométriques des échelles ORS et SRS entre la version originale anglophone et la version française des échelles ; avec une structure factorielle comparable, une même fidélité des scores de la mesure, une même validité de contenu, une même validité de construit, et une même validité discriminante ou de sensibilité au changement calculées à partir des scores obtenus aux versions anglophones et francophones de ces échelles. Nous nous attendons notamment à ce que les scores obtenus aux échelles ORS et SRS soient corrélés significativement et positivement avec les scores obtenus avec des échelles de référence francophones mesurant leurs construits associés respectifs ; et que les scores obtenus à l'échelle ORS puissent être significativement corrélés avec les scores

obtenus avec une autre échelle pour laquelle le construit mesuré est considéré comme théoriquement lié.

En outre et en regard des résultats des adaptations étrangères précédentes (en allemand et en espagnol), nous faisons l'hypothèse que les données prénormatives vont différer, et que les indices de changement fiables et de seuil cliniques vont différer de ceux de la langue originale. Nous opérationnaliserons plus avant ces hypothèses lors de la validation française de ces échelles qui sera réalisée à la section 2.1.

1.4.3. Hypothèses relatives aux liens entre le feedback non verbal et le feedback instrumental dans la pratique ROM.

Afin d'apporter des premiers éléments, qui nous l'espérons, seront susceptibles de mieux expliquer les effets du feedback instrumental en psychothérapie, et de soumettre en partie à l'épreuve des faits les modèles palliatifs et supplétifs du feedback instrumental, nous souhaitons tester les habiletés de psychothérapeutes qui sont engagés dans une pratique ROM dans la reconnaissance des signaux non verbaux.

Nous postulons que les meilleurs thérapeutes en condition de feedback instrumental ROM seront également les thérapeutes les plus performants dans la perception des signaux non verbaux émotionnels (macro et micro, vocaux, faciaux et corporels).

Plus précisément, nous postulons que les thérapeutes les plus efficaces en condition de feedback instrumental seront les plus performants dans la détection des expressions émotionnelles négatives (macro et micro) en comparaison avec la reconnaissance des émotions négatives.

Nous postulons également et que l'habileté dans la détection des micro-émotions micro, en particulier négatives, sera associée à de meilleurs résultats en comparaison avec l'habileté dans la détection des macro-émotions. Nous opérationnaliserons plus avant ces hypothèses lors

de notre étude pilote d'évaluation des capacités de reconnaissance des psychothérapeutes en condition de feedback instrumental qui sera présentée à la section 2.2.

2. PARTIE EMPIRIQUE

1.5. Étude n°1 : Validation française des Échelles de Résultats et de l'Alliance thérapeutique ORS et SRS (Miller et al., 2005) ; l'étude VFERA

1.5.1. Origines de la démarche

La méthodologie du Routine Outcome Monitoring (ROM) a été développée, nous l'avons vu pour répondre aux enjeux des mauvais résultats dus aux traitements, et aux abandons en cours de soins. Notamment pour aider les cliniciens à mieux identifier les patients à risque de se détériorer et à mieux adapter le traitement délivré (Lambert, 2015). Ceci implique de pouvoir suivre de séance en séance l'évolution du patient (les résultats c'est-à-dire le changement thérapeutique), mais également en fonction des systèmes de ROM, l'évolution de l'alliance thérapeutique, afin d'adapter la thérapie et de mieux engager les patients dans les soins.

De nombreux systèmes basés sur le feedback des patients ont été développés et sont utilisés à travers le monde, mais à notre connaissance, le ROM est moins fréquemment utilisé dans les pays francophones et il n'existe pas de système complet et suffisamment simple - demandant peu de temps d'utilisation pour pouvoir être directement implémenté en pratique clinique courante - et qui soit à la fois reconnu comme validé empiriquement (notamment par les autorités de santé aux USA) et qui ait été adapté et validé avec la population française.

De ce fait, le système PCOMS (the Partner for Change Outcome Management System, PCOMS ; Miller et al., 2005), représente une option intéressante puisque ce système est à la fois reconnu comme empiriquement validé et enregistré au répertoire américain des programmes et pratiques empiriquement validés (répertoire NREPP de la SAMSHA) ¹¹, et que son utilisation est documentée comme plus aisée et moins chronophage que des échelles, certes

¹¹ <http://www.nrepp.samhsa.gov/ViewIntervention.aspx?id=249>

plus complètes, mais aussi sujettes à être rapidement abandonnées par les cliniciens . Dans les recommandations de la revue Cochrane sur l'utilisation de la pratique ROM pour le traitement des troubles mentaux courants (Kendrick et al., 2016), les auteurs recommandent d'ailleurs de considérer la rapidité d'administration des échelles de feedback dans le choix d'utiliser un système ROM et évoquent en ce sens l'utilisation de l'Échelle d'Évaluation des Résultats de la Séance (Outcome Rating Scale, ORS ; Miller et al., 2005) qui est partie intégrante du système PCOMS.

Le système PCOMS comprend donc, nous l'avons vu, une échelle de mesure des résultats psychothérapeutiques, l'échelle ORS (annexe 2) et une échelle de mesure de l'alliance thérapeutique l'échelle SRS (Session Rating Scale ; Miller et al., 2005 ; annexe 3). Le processus clinique d'administration continue de ces deux échelles pour suivre les résultats en continu et améliorer le traitement est appelé Traitement Orienté par le Feedback (Feedback Informed Treatment ou FIT).

L'Échelle d'Évaluation des Résultats, ORS (Outcome Rating Scale ; Miller et al., 2003) est une échelle très brève et auto-administrée, destinée à mesurer les dimensions du bien-être personnel qui sont supposées manifester des changements positifs quand un traitement psychothérapeutique est efficace. Ce questionnaire a directement été inspiré des travaux de Michael Lambert sur le Routine Outcome Monitoring du changement thérapeutique et par la construction du questionnaire OQ-45 qui mesure la détresse psychologique sur trois dimensions (une dimension de détresse psychologique subjective, une dimension de relations interpersonnelles et une dimension de relations sociale).

Le questionnaire ORS comprend quatre échelles visuelles analogiques (EVA) qui mesurent 10 cm chacune dans la présentation au format papier du questionnaire. Chaque échelle mesure le niveau de bien-être sur chaque dimension, avec l'extrémité gauche de l'échelle associée au plus bas niveau (0) et l'extrémité droite de l'échelle associée au plus haut niveau (10). Les 4 dimensions mesurées sont respectivement la dimension de bien-être individuel (item

a), de bien être interpersonnel (item b), de bien-être social (item c), et de bien-être global (item d).

Les questionnaires qui sont chronophages ont moins de chance d'être réellement administrés en continu en pratique clinique courante (Brown, Dreis, & Nace, 1999 ; Miller et al., 2003; Seidel et al., 2016) et l' échelle ORS a été dans ce sens développée comme une alternative aux questionnaires de mesure du changement thérapeutique qui sont plus long à administrer, tout en permettant d'être utilisé dans une perspective de soins basée sur la mesure (« measurement based care ») du fait de qualités psychométriques satisfaisantes documentées.

Le score de chaque item étant compris entre 0 et 10, le score total obtenu à l'échelle ORS est obtenu par addition des scores obtenus aux items a, b, c et d. Il est compris entre 0 et 40. De plus hauts scores à l'item ou au questionnaire indiquent un degré supérieur de bien-être.

Dans le processus de thérapie guidée par le feedback, le questionnaire ORS est administré à chaque début de séance pour servir de mesure des résultats (du changement thérapeutique).

L'Échelle d'Évaluation des Résultats de la Séance ou SRS (Session Rating Scale 3.0 ; Miller et al., 2003) est également une échelle ultra-brève et auto-administrée, qui est destinée à mesurer les trois dimensions de l'alliance thérapeutique correspondantes au modèle de Bordin (1979). Ce questionnaire est également constitué de quatre échelles visuelles analogiques mesurant 10 cm dans la présentation au format papier. Les items mesurent respectivement l'entente/la relation entre le patient et le thérapeute (item a), l'entente sur les buts de la thérapie (item b), l'entente sur les taches thérapeutiques (item c), et une quatrième question est destinée à mesurer la qualité globale de la séance (item d).

Pour chaque dimension, l'extrémité gauche de l'échelle correspond au niveau le plus bas (0) et l'extrémité droite au niveau le plus haut (10). Les scores obtenus à chaque item étant compris entre 0 et 10, le score total à l'échelle, obtenu par l'addition des quatre items, est

compris entre 0 et 40. Un score plus haut à l'item ou à l'échelle est associé à une meilleure alliance thérapeutique.

Les échelles d'évaluation du changement thérapeutique (ORS) et d'évaluation de l'alliance thérapeutique (SRS) ont été initialement traduites en français, mais n'ont pas fait l'objet d'un travail de validation transculturelle. Ce travail est pourtant nécessaire, car si ces échelles peuvent être directement utilisées pour favoriser le recueil du feedback du patient en condition clinique, l'étude des qualités psychométriques de la version française s'impose quand il s'agit d'apporter des données probantes susceptibles d'orienter les décisions affectant par exemple la politique d'encadrement des soins.

Dans le but de documenter la légitimité et l'applicabilité de ces échelles, pour les populations francophones, et après avoir recherché d'éventuels collaborateurs possédant des données normatives ou intéressés par un travail de validation, nous avons consulté l'auteur principal des mesures PCOMS, le Dr Scott D Miller et nous avons initié en mars 2013, un projet de validation unifiée des échelles avec différents cliniciens francophones.

Documenter l'utilité clinique et donc sociale de l'utilisation des échelles renvoyait à interroger les qualités psychométriques de ces échelles dans leur versions francophones (validité de structure, analyses de la validité et de la fidélité) ainsi qu'à déterminer les indices statistiques de seuils cliniques et de changement fiables relatifs à ces échelles. Il est également apparu nécessaire de documenter les premières données normatives issues de populations francophones obtenues auprès d'un échantillon de patients suivis en psychothérapie (population clinique), mais également recueillies avec un échantillon de sujets tout venant ne recevant ni traitement psychothérapeutique, ni traitement médicamenteux.

1.5.2. Présentation de la méthodologie générale d'adaptation transculturelle.

1.5.2.1. Projet de départ

Nous avons développé un projet de validation et de collecte des données en nous appuyant sur la méthodologie générale de validation transculturelle des échelles proposée par Valleyrand (1989), que nous avons complété par les recommandations de la commission internationale des tests, notamment en ce qui concerne la standardisation de notre protocole (ICT, 2016). Nous avons tout d'abord étudié la méthodologie des études d'adaptation des échelles ORS et SRS, dans d'autres langues étrangères, pour inscrire notre travail dans la continuité de la démarche empirique de nos collègues, mais également pour essayer de dépasser des potentielles limitations méthodologiques dues aux précédentes démarches de validation. Nous avons repris contact avec l'auteur des échelles, le Dr Scott D Miller, pour recueillir son agrément quant au processus de validation, le système PCOMS étant un système propriétaire, soumis à une licence d'utilisation, et également pour recueillir son autorisation d'adapter et d'administrer les échelles ORS et SRS à l'aide d'un support informatique (ordinateur ou tablette). Nous l'avons convié à participer au processus de validation, notamment en lui demandant de prendre part aux discussions d'un premier comité de révision de la traduction initiale des échelles ; puis au long cours à chaque fois qu'une question méthodologique impliquait sa participation.

Nous avons présenté ce projet et les étapes de validation transculturelle à des collègues francophones (Canada, Belgique), avec en première intention, la possibilité d'initier une validation des versions francophones unifiées des deux questionnaires, aussi bien en Europe francophone qu'outre atlantique. La validation des échelles ORS et SRS en français débuta en 2013, quand nous avons finalement coordonné un comité technique composé de cliniciens francophones : avec les cliniciens de l'agence Valoris pour les enfants et les Adultes de Prescott-Russel (Valoris for Children and Adults of Prescott-Russell Agency au Canada), Raymond Lemay, Chantal Tassé, Madeleine Lalonde, et Nicholas Cardinal, et avec John

Deltour (pour la Belgique) ; dans le but d'arriver à un consensus sur la traduction des échelles, afin d'assurer la comparabilité des données entre les différentes recherches menées dans les différents groupe francophones et planifier de futures récoltes de données et analyses.

1.5.2.2. Le processus d'adaptation transculturelle

Nous avons distingué deux grandes étapes dans la séquence de validation de la méthodologie dégagée par Valleyrand (1989) : une étape de production ou d'adaptation, dans le but de rendre une version dite expérimentale des échelles, organisée ici, autour de la révision de la traduction initiale des échelles. Cette étape doit permettre de réviser et d'unifier toutes les différentes versions des échelles PCOMS, y compris l'ORS et la SRS. Une seconde étape d'administration à proprement parler des échelles ou de validation auprès de populations cliniques et non cliniques afin d'effectuer tous les traitements statistiques appropriés pour évaluer les qualités psychométriques. Ces deux grandes étapes peuvent être détaillées selon la séquence suivante.

Étape 1 : Une préparation de la version dite préliminaire des échelles ORS et SRS.

Une révision par le comité de la traduction initiale des échelles PCOMS avec un examen indépendant des trois groupes (Canada, Belgique, France), effectué par les membres du comité.

Étape 2 : Une évaluation de la version préliminaire.

Le comité évalue l'adéquation de la version française (VF), c'est-à-dire la conformité des consignes et de la formulation des items traduits par rapport à la version originale (VO ; consigne et items sources) et vérifie la formulation technique des questionnaires. Il aboutit à une formulation consensuelle tenant compte si possible de l'adéquation aux contextes culturels nord-américains (Canada francophone) et Européens. Cette version est dénommée version expérimentale puisqu'elle peut faire l'objet d'étude. Ces deux étapes sont réalisées par le comité.

Étape 3 : Un pré-test de la version expérimentale française.

Un échantillon de la population cible, c'est-à-dire des personnes dont l'âge se situe entre 18 et 70 ans, issues de la population française, et qui sont de ce fait potentiellement susceptibles de recevoir une psychothérapie à un moment donné de leur vie, et donc de donner leur feedback sur le processus des soins et les résultats dus aux traitements, évaluent la clarté de la formulation des consignes et des items. L'évaluation de la clarté se fait sur une échelle de 0 à 7 (selon les recommandations de Valleyrand, 1989). S'il y a ambiguïté, il s'agit de recueillir le feedback du sondé. Si les scores moyens aux consignes et aux items sont supérieurs à 4, cette version est retenue en tant que version expérimentale française. Autrement, il y a lieu de réviser l'échelle pour améliorer la formulation.

Étape 4 : Évaluation de la validité convergente

La validité convergente ou « concomitante » (de l'anglais concurrent validity), fait référence à la capacité d'un instrument de mesure à évaluer un construit spécifique. Elle peut être documentée quand les mesures recueillies sur deux échelles censées évaluer le même construit convergent. Selon Valleyrand (1989), cette validité peut être démontrée empiriquement quand les scores obtenus à un test (par exemple ici les scores obtenus à l'échelle ORS) sont corrélés significativement avec un autre test (de référence pour le construit questionné ; ici pour l'ORS une mesure de la détresse psychologique).

Il s'agit donc d'administrer concomitamment la version française des échelles que nous souhaitons valider et la version française d'échelles françaises validées mesurant le même construit ; ce qui revient à d'administrer à la même occasion de mesurage la version française de l'ORS ainsi que la version française d'une échelle validée de mesure de la détresse psychologique utilisable pour suivre le changement psychothérapeutique en cours de traitement. Et également d'administrer à la même occasion de mesurage, la version française de l'échelle SRS ainsi que la version française d'une échelle validée mesurant l'alliance thérapeutique.

La validité de contenu fait référence à la capacité des items qui constituent l'échelle de mesure d'évaluer le construit pour lequel ils ont été conçus – cette capacité peut être mise en évidence lors de la construction du questionnaire en impliquant des experts du domaine étudié, mais aussi par l'étude des corrélations entre les scores obtenus à l'échelle qui est analysée et les scores obtenus à une échelle dont la validité de contenu est connue dans la langue cible.¹²

En étudiant la matrice de corrélation des scores, si les tests sont concluants, nous pourrions affirmer que la validité de contenu et la validité convergente sont conférées à la version française des échelles ORS et SRS (selon la méthodologie de Valleyrand, 1989). Si les tests s'avèrent non significatifs, une révision des échelles s'imposera avant une nouvelle administration à d'autres participants.

Etape 5 : Analyse de la fidélité des scores de la mesure

La fidélité des scores de la mesure fait référence à la constance dans l'évaluation d'un construit. Un score de mesure est réputée fiable, et donc précis, s'il mesure toujours de manière similaire un même construit (Valleyrand, 1989) ; c'est-à-dire s'il renvoie des données constantes en mesurant le même niveau du construit à différentes occasions. Plusieurs indices permettent d'asseoir la fidélité des scores d'une mesure. Dans le cas de l'exploration d'un construit mesuré sur une échelle continue, le calcul de la consistance interne, c'est à dire le calcul de l'homogénéité (la corrélation entre les scores obtenus aux différents items en tenant compte du nombre des items) correspond au coefficient alpha de Cronbach.

Dans notre étude, il s'agit d'utiliser les scores obtenus aux items de l'échelle ORS et SRS pour calculer ce coefficient.

¹² <https://www.irdp.ch/institut/validite-contenu-2181.html>

Étape n° 6 : Analyse de la validité de construit, et de la sensibilité au changement.

La validité de construit renvoie au fait que l'échelle mesure le construit tel qu'il a été défini théoriquement : au niveau de la structure du construit et des éventuels sous dimensions, ainsi qu'au niveau des relations entre le construit et les corrélats du construit (Vallerand, 1989).

Le premier niveau d'analyse de la validité de construit concerne donc l'analyse des items et de la structure factorielle. Pour des échelles dont la structure factorielle a été mise en évidence dans la langue d'origine, il s'agit de procéder à une analyse factorielle confirmatoire (AFC). Cette analyse doit permettre de tester si les échelles possèdent en français la même structure factorielle que leurs équivalents formulés en langue anglaise. Dans le cas d'échelles multidimensionnelles, il convient d'effectuer une analyse des relations entre les différentes dimensions.

Enfin, le test des corrélats du construit psychologique avec l'échelle qui a été adaptée en français, amène des arguments complémentaires au fait que l'échelle mesure bien le construit qui est théoriquement en relation avec ces autres construits. Dans cette étude, il est attendu que l'ORS soit significativement corrélée avec les scores d'une échelle mesurant un construit qui est censé être relié à la mesure du bien-être ou de manière inversée la détresse psychologique. Au niveau de l'échelle SRS, il est également attendu que l'alliance prédise de meilleurs résultats au traitement et donc que les scores obtenus à l'échelle SRS prédisent les scores subséquents obtenus à l'échelle ORS.

La sensibilité au changement, la capacité de l'échelle à mesurer une variation d'un état, peut être établie, en étudiant la significativité des différences de scores qui sont théoriquement attendues et mesurées par l'échelle adaptée en français. Ici l'échelle de mesure du changement adaptée en français (ORS), doit refléter une différence significative, entre une première mesure à l'entrée dans le traitement pour le groupe clinique, et la dernière séance. Pour le groupe témoin (non clinique), l'étude des différences de scores entre le test et le re-test ne doit pas révéler de différence statistiquement significative.

Dans le cas où ces différentes analyses ne seraient pas concluantes, l'échelle ferait l'objet d'une révision avant une nouvelle administration.

Étape 7 : Obtention des données normatives préliminaires.

La documentation des moyennes et des écart-types, obtenues à partir des scores recueillis sur les échelles ORS et SRS, dans les échantillons cliniques et/ou non cliniques constituent des données prénormatives de référence pour la population clinique considérée. Ces résultats doivent être comparés à ceux obtenus par les sujets qui ont reçu l'administration des échelles dans leurs versions originales, et dans d'autres adaptations étrangères, car elles reflètent l'adaptation au contexte culturel.

Il s'agit également de s'assurer que ces différences soient bien significatives, ce qui constitue un argument supplémentaire quant au potentiel de l'échelle (ORS) à discriminer les différentes populations (Cliniques et Non Cliniques).

L'enjeu de cette validation est notamment de tenir compte des possibles limitations méthodologiques des études antérieures pour tenter d'amener un éclairage supplémentaire sur les qualités psychométriques de ces échelles. Afin d'étudier les qualités psychométriques des échelles ORS et SRS, nous avons recherché des échelles de référence pour les construits mesurés par les échelles ORS et SRS, validées en français, et soit libres d'utilisation, soit utilisables dans une recherche après accord de l'auteur, afin d'analyser les liens (différences, convergences) entre les scores obtenus aux échelles ORS et SRS.

1.5.3. Mesures utilisées

Outre les versions françaises des questionnaires ORS et SRS notre protocole contenait les mesures suivantes.

1.5.3.1. Le questionnaire d'impact MI-22 (Outcome Questionnaire-22, OQ-HAI-22, Lambert & Burlingame, 1996b, version française validée par Flynn et al., 2003).

Le questionnaire MI-22 est une échelle auto-administrée composée de 22 items, conçue pour évaluer le degré de détresse psychologique expérimentée par les patients au cours de leur thérapie et de ce fait le degré de changement relié au traitement. La MI-22 est une échelle basée sur un échantillon réduit des items du questionnaire de changement OQ-45.2 qui est considéré comme un « étalon-or » de la mesure des résultats psychothérapeutiques (OQ-45.2, Lambert & Burlingame, 1996a). Les items mesurent la détresse expérimentée dans trois dimensions cliniques (la détresse subjective, les relations interpersonnelles, et le rôle social). Chaque item correspond à une réponse en 5 points sur une échelle de Likert avec des scores de 0 à 4. Le score total correspond à l'addition du score des trois domaines de fonctionnement et est compris entre 0 et 88. Les plus hauts scores indiquent un plus haut niveau de détresse sur chaque domaine mais aussi sur le score total. Nous avons choisi d'utiliser la version française du questionnaire OQ-22 pour examiner la validité convergente des scores obtenus à cette échelle avec ceux obtenus lors de l'administration de l'échelle ORS. Nous avons obtenu l'autorisation du Dr Michael Lambert de l'utiliser directement dans cette recherche et de l'adapter pour une administration informatisée. La version utilisée a été obtenue auprès des auteurs de la validation française.

1.5.3.2. L'échelle Working Alliance Inventory-Short form (WAIS-S, Tracey and Kokotovic, 1989 ; version française validée par Corbière et al. (2006)).

Il s'agit d'une courte échelle auto-administrée, composée de 12 items, et conçue pour évaluer la qualité de la relation collaborative entre le thérapeute et son patient, selon le modèle de Bordin (1979). L'échelle WAI-S est composée de questions évaluant les trois dimensions cliniques de l'alliance de travail (la relation, l'entente sur les buts, l'entente sur les tâches). Chaque réponse suit une échelle de Likert en sept points. Le score total est obtenu par la somme des scores aux trois dimensions (compris entre 0 et 84). Les scores les plus hauts sont la manifestation d'une meilleure alliance. Nous avons choisi d'utiliser la version française de la WAI-S pour explorer la validité convergente des scores avec les scores obtenus à l'échelle SRS, après avoir remplacé le terme "mon intervenant en emploi" par le terme « mon thérapeute » conformément aux recommandations du Dr Corbières qui nous a autorisé à utiliser cette échelle au format électronique pour cette recherche. L'échelle a été validée dans un contexte psychothérapeutique. Nous avons légèrement modifié la formulation des items 5 et 12 pour mieux coller à une formulation française.

1.5.3.3. L'échelle de Kessler-6 ou K6 (Kessler-6 Scale; Kessler and al., 2003; National population Health survey, 1999 ; étude de validité française réalisée par Arnaud et al., 2010).

Il s'agit d'un questionnaire ultra-bref auto-administré destiné à évaluer le degré de détresse psychologique générale présenté par un individu.

Ce questionnaire est composé de 6 items dont le format de réponse suit une échelle de Likert en 5 points. Le score total est obtenu par la somme des scores aux 6 questions et il est compris entre 0 et 24. De plus hauts scores indiquent de plus hauts scores de détresse psychologique. L'échelle K6 a été utilisée dans des enquêtes mondiales de L'OMS comme un outil de dépistage de la détresse psychologique et dans des publications françaises (Nguyen et al., 2012). Nous avons obtenu l'autorisation du Professeur Ronald Kessler d'utiliser la version

traduite officielle de cette échelle et de modifier la consigne temporelle pour coller à celle de l'administration de l'ORS (7 jours). Nous avons choisi d'utiliser l'échelle K6 pour examiner la validité convergente des scores obtenus avec ceux de l'échelle ORS.

1.5.3.4. Le Big Five Inventory français - The Big Five Inventory (BFI, John, Donahue & Kentle, 1991 ; version française validée par Plaisant et al., 2010)

Ce questionnaire auto-administré a été conçu pour évaluer les traits de personnalité selon les cinq dimensions du modèle des "5 grands" (L'Extraversion, l'Agréabilité ou Aménité, le caractère consciencieux, L'affectivité négative, et l'ouverture). L'échelle est composée de 45 items suivant une réponse de style Likert sur un format en 5 points. Les plus hauts scores sur une dimension indiquent une présence plus marquée de la dimension. Nous avons choisi d'utiliser la version française de la BFI pour évaluer la validité de critère de l'échelle ORS

La figure 6 reprend le recueil des informations utiles selon la séquence de Valleyrand

Validation française des échelles PCOMS (ORS/SRS)

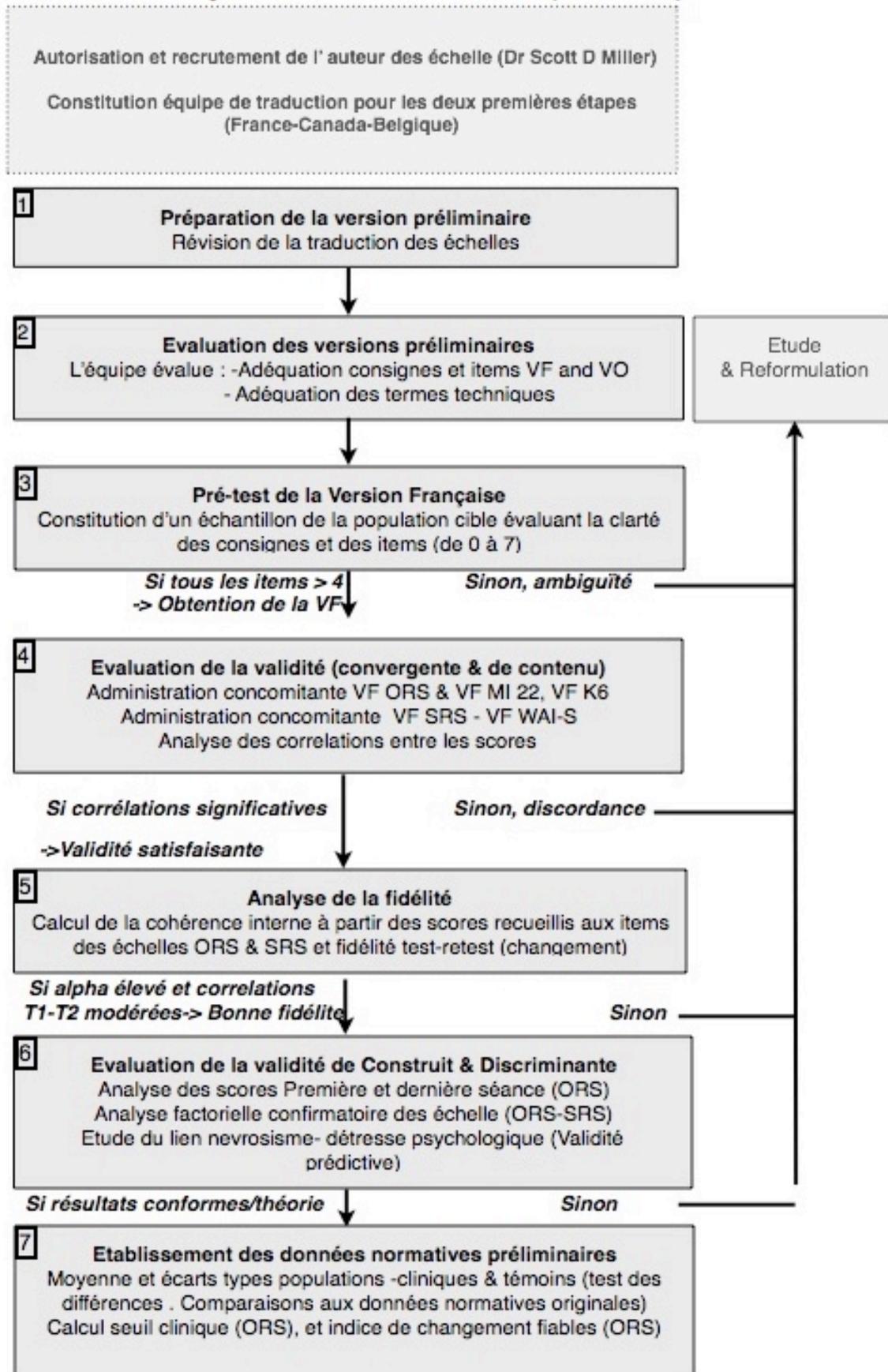


Figure 6. Séquence d'adaptation transculturelle (inspiré de Valleyrand, 1989)

A des fins de contrôles, il est également nécessaire de recueillir des variables contextuelles : l'âge, le sexe, le statut marital, le statut professionnel, et le niveau d'éducation. Le recueil des données dans notre protocole est précisé à l'annexe 4.

Les deux premières étapes du processus d'adaptation transculturelle ont permis de créer une version unifiée des échelles validée par le Dr Scott D Miller. La formulation de ces versions a également été vérifiée, après production, par le Professeur Kamel Gana.

1.5.3.5. Étape de pré-test des échelles.

Un pré-test de clarté et de compréhension de la version expérimentale de ces échelles a finalement été administré à 20 personnes, composant un échantillon hétérogène extrait de la population française, et qui pourraient être susceptibles de recevoir un traitement psychothérapeutique.

Lors de ce pré-test, le contexte minimal d'administration des échelles a été spécifié et un effort d'imagination a été demandé aux participants : l'échelle ORS a été présentée comme une échelle mesurant le bien-être et le changement en cours de thérapie ; l'échelle SRS a été présentée comme une échelle visant à recueillir du feedback sur la relation entre un thérapeute et son patient.

Les versions françaises des échelles ORS et SRS ont été bien reçues par les participants avec des scores moyens aux instructions et aux items tous supérieurs au seuil critique de clarté de 4, défini par Valleyrand (1989). Le score moyen à l'item (d) de l'ORS (Globalement, perception globale du bien-être) a été coté à un niveau inférieur en comparaison des autres items ce qui laisse à penser que les participants ont eu besoin de plus de contexte pour distinguer cet item d de l'item a (Bien être individuel) ; dans la démarche FIT, il est en effet nécessaire de donner un maximum d'informations sur les échelles aux patients lors de la première passation et il est admis qu'il puisse être nécessaire de leur redonner des informations complémentaires sur l'administration de l'échelle à chaque fois que cela s'avère nécessaire.

1.5.4. Procédure et Participants

1.5.4.1. Procédure Générale

Nous avons utilisé la plateforme Limesurvey (Carsten Schmitz, 2003) du laboratoire de psychologie de l'Université de Bordeaux (EA4139) pour créer nos questionnaires et les administrer en ligne. Cette interface, permet d'accéder aux questions de manière anonyme (avec un accès par code) et de manière sécurisée.

Nous avons dû modifier la feuille de style disponible pour l'affichage des questionnaires afin de coller au plus près de la présentation papier des questionnaires ORS et SRS, Limesurvey ne permettant pas nativement de créer des échelles visuelles analogiques avec du texte disposé au-dessus des échelles et avec un espacement convenable des multiples EVA. Un exemple de l'affichage du premier item de l'échelle SRS est affiché ci-dessous (Figure 7)



Figure 7 : Item a de l'échelle SRS (étude VFERA)

Nous avons créé une brochure d'administration à destination des thérapeutes (annexe 8) qui leur expliquait comment accéder aux questionnaires pour ensuite laisser les patients y répondre seuls, que faire en cas de difficultés techniques, et comment utiliser la fiche de pointage (annexe 9) pour obtenir des informations complémentaires sur le traitement. Les participants issus d'une population non clinique ont également reçu une brochure d'auto-administration leur expliquant comment participer et que faire en cas de problème technique (brochure en annexe 10). Nous avons bêta-testé auprès de plusieurs utilisateurs cette plateforme et pratiqué des « crash test » de l'application pour nous assurer de la facilité d'utilisation et de récupération des informations en cas de problèmes techniques.

Nous avons rencontré des problématiques techniques au cours de notre étude (panne de serveur, panne du navigateur, enregistrements vides, oubli de code, oublis de séance chez le thérapeute) et seules les données complètes recueillies ont été étudiées. Nous avons effectué un très important travail de consolidation des données de manière manuelle à la fin de l'extraction depuis la base Limersurvey. Nous avons notamment rétabli la séquence des séances, quand une séance n'avait pas été administrée par le thérapeute et que le décompte de séance était faux pour cette raison, en analysant l'horodatage dans le logiciel.

1.5.4.2. Participants

Échantillons cliniques

Nous avons constitué deux échantillons cliniques. *Un premier échantillon clinique* composé de 28 patients entamant leur première séance de psychothérapie et recrutés à partir de la patientèle de 4 psychologues (3 femmes et un homme) qui ont répondu positivement à notre appel de participation à cette étude. Initialement, 7 thérapeutes s'étaient montrés intéressés par cette étude et déclarés prêts à participer.

Les cliniciens ont signé un consentement éclairé écrit leur expliquant les modalités de la recherche (Annexe 5), qu'ils n'auraient pas connaissances des résultats de leurs patients (les plaçant en condition clinique non feedback ou CNF) et que l'investigateur de l'étude n'aurait pas accès à l'identité des patients. Les clients de ces 4 psychothérapeutes ont signé un consentement éclairé électronique leur précisant les objectifs et modalités de cette recherche (Annexe 6), le fait que les données seraient recueillies de manière anonyme et que leurs thérapeutes n'auraient pas accès aux données, ce qui est à même de réduire les biais de désirabilité sociale et de favoriser un feedback plus libre.

4 thérapeutes (3 femmes, 1 homme) récoltèrent les données de 28 patients avec un échantillon de patients sondés compris entre 1 et 16 patients selon les thérapeutes et un âge des patients compris entre 20 et 45 ans ($M = 30.43$, $ET = 6$).

Au niveau des données sociodémographiques de contrôle, dans ce groupe, 71.4% des patients ont déclaré être en couple (mariés ou en cohabitation) et 28.6% n'étaient pas au moment de l'étude dans une relation de couple (célibataire, divorcés ou veuf/ve). En ce qui concerne le niveau d'éducation 96.4% des participants de ce groupe se situaient au niveau du baccalauréat français ou à un niveau supérieur. Pour l'appartenance à une catégorie socioprofessionnelle, 85.7% reportaient être en activité.

3 des 4 thérapeutes ont rempli la fiche de pointage relative aux soins distribués à chaque patient et reprenant des informations sur leur pratique et leur expérience. L'âge moyen des thérapeutes était de 52 ans ($M = 52.33$, $ET = 5.86$), leur expérience clinique moyenne en années était de 8 ans ($M = 8.33$, $ET = 2.08$), leur approche de traitement utilisée la plus représentée était la thérapie comportementale et cognitive. Pour leurs 24 clients respectifs, le nombre de séances a été compris entre 1 et 9 chez leur psychothérapeute. 16.7% des patients ont été considérés comme "mandaté à recevoir des soins" parce qu'ils ne se sont pas rendus en thérapie de leur propre chef. 79.2% ($n = 19$) des patients consultaient pour des troubles anxieux ou des troubles de l'humeur avec une sous-proportion de 4.2% des patients qui présentaient concomitamment de l'anxiété et des problèmes d'abus de substances, 20.8% des patients sont venus consulter pour d'autres raisons incluant des problématiques de santé des problèmes somatiques ou des troubles de la personnalité. 37.5% des patients recevaient un traitement médicamenteux.

Les thérapeutes ont reçu des codes d'accès spécifiques pour chacun de leurs patients. Les thérapeutes avaient pour consigne d'arrêter d'enregistrer les données des patients dans les situations suivantes :

- Si le patient revenait en thérapie après un délais supérieur à 8 semaines depuis la dernière séance. Il s'agit d'un critère purement empirique qui est utilisé dans la pratique Feedback Informed Treatment (FIT) pour qualifier un nouvel épisode de traitement. En effet, la mesure enregistrée sur une séance n'apparaîtrait plus prédictive de la prochaine

séance au-delà de ce délai de 8 semaines et comme les algorithmes prédictifs ont pour vocation d'estimer la trajectoire du changement, la mesure ne serait plus fiable après ce délai. Nous avons donc décidé de limiter les enregistrements successifs de séance à cette durée afin de pouvoir tester des hypothèses appropriées.

- Si la thérapie était terminée.
- Si le patient avait abandonné la thérapie.
- Si le patient décidait de ne plus enregistrer ses séances.

Sur l'échantillon de 24 patients pour lesquels les thérapeutes ont fournis des renseignements complémentaires, l'enregistrement fut stoppé pour un délai supérieur entre deux séances consécutives à 8 semaines, dans 12.5% des cas, et les patients ont abandonné leur traitement dans 37.50% des cas ; les patients restant étaient soit toujours en traitement à la fin de la collecte des données (25%) ou avaient terminé leur thérapie (25 %). Ces données sont importantes car elles sont assez similaires aux taux d'abandons rapportés dans la littérature.

Un second échantillon clinique (n = 28) a été constitué de patients que nous avons reçu en psychothérapie en consultation externe, et qui ont commencé leur premier épisode de traitement, en utilisant les échelles PCOMS dans leur nouvelle version unifiée.

Ces patients ont rempli un formulaire de consentement éclairé les informant que leur thérapeute utilisait des échelles de feedback pour suivre leurs progrès en termes de résultats et d'alliance thérapeutique. Dans ce formulaire, ils autorisaient expressément l'utilisation de leurs données à des fins de recherche. Nous avons administré l'ORS et la SRS à chaque séance de thérapie et utilisé le feedback des patients pour adapter le traitement, comme nous procédons de manière habituelle, et ces patients ont été de ce fait en condition feedback (CF). Nous avons évalué les difficultés des participants de manière à produire une analyse fonctionnelle de leur problématique et de ce fait, seulement les informations nécessaires au traitement rapportés par les patients sont reproduites ici pour donner les informations sur le contexte des soins ; pour ces

raisons elles n'ont pas été catégorisées selon la même séquence que dans le groupe CNF. Dans ce second groupe clinique le mode de traitement privilégié était la thérapie comportementale et cognitive (TCC). Les séances disponibles pour l'analyse variaient de 1 à 15 sachant que seules les séances issues d'un premier épisode de traitement, avec un délai entre les séances ne dépassant pas 8 semaines (2 mois ou de manière moins conservatrice une durée de 61 jours) ont été analysées.

Au niveau des variables sociodémographiques, les patients du groupe CF étaient âgés de 18 à 71 ans ($M = 34.61$, $ET = 12.89$), 71.4% ont déclarés être en couples. 53.6% exerçaient une activité professionnelle contre 46.4% qui n'en exerçaient pas (qui suivaient des études, n'occupaient pas d'emploi ou étaient à la retraite).

57.2% des patients consultaient pour des problèmes d'humeur, d'anxiété ou des problèmes émotionnels, 21.4% recevaient des soins pour des problèmes relationnels, 7.1% pour des problèmes professionnels, et 14.3% pour d'autres raisons (incluant des problèmes de stress, des conduites addictives, ou d'autres problématiques de santé). 42.9% des patients étaient identifiés comme recevant un traitement médicamenteux à un moment donné de la thérapie.

Dans ce sens les patients étaient assez représentatifs des problématiques les plus couramment rencontrées pour des soins reçus en consultation externe.

Ces deux échantillons cliniques vont être analysés de manière comparative avec un échantillon non clinique dans le but d'évaluer les propriétés discriminantes de l'échelle ORS (sensibilité au changement). Les propriétés psychométriques pour les échelles ORS et SRS, vont également être analysées pour les échantillons cliniques et non cliniques, ce qui permettra de documenter des données pré-normatives.

1.5.4.3. Critères d'inclusion et d'exclusion

Les patients devaient être âgés de plus de 18 ans, et devaient participer à leur première séance de psychothérapie avec leur psychologue. La première séance de psychothérapie étant

ici définie comme la toute première séance avec le psychologue, même si leur praticien adoptait un format d'évaluation, lors de cette séance, et de recueil des informations cliniques importantes (pour une analyse fonctionnelle) plutôt qu'une interaction psychothérapeutique à proprement parler.

Les patients qui présentaient des troubles ou un état de crise jugé incompatible par le clinicien avec l'administration des échelles ORS et SRS étaient exclus du protocole (par exemple un trouble psychotique, un état de crise, des difficultés marquées dans la compréhension de la langue française ou dans l'administration autonome des mesures du fait d'un handicap cognitif ou de problèmes moteurs). Ces conditions ont été posées pour coller au plus près aux protocoles de validation initiaux de ces échelles.

1.5.4.4. Échantillons issus de la population française

Pour pouvoir effectuer des comparaisons de scores entre un échantillon non clinique issu de la population française et la population clinique de patients, nous avons recruté des étudiants de deux universités françaises (Bordeaux et Paris). Ces étudiants ont reçu une brochure explicative et un code d'accès à notre enquête par e-mail. Les volontaires ont rempli un questionnaire en ligne lors d'une première occasion (test) et lors d'une seconde occasion re-test sur leur ordinateur ou tablette. 154 étudiants ont rempli le questionnaire à la première occasion et n à la seconde occasion. Ces étudiants ont été pleinement informés des finalités de la recherche et ont signé un consentement éclairé électronique (annexe 7).

Parmi ces 154 étudiants, 24 ont rapportés qu'ils recevaient actuellement des soins et/ ou un traitement psychotrope et pour cette raison nous les classerons dans l'échantillon étudiant Clinique (EC) et 130 ne recevaient pas de soins ou de traitements médicamenteux pour des problématiques psychologiques et pour cette raison nous les classons dans un échantillon étudiant non clinique (ENC). L'échantillon EC ne sera utilisé que pour des analyses complémentaires.

L'échantillon ENC était composé de 83.8% de femmes et de 16.2% d'hommes, l'échantillon EC était lui constitué de 91.7% de femmes et 8.3%. Les étudiants étaient âgés de 18 à 59 ans dans l'échantillon ENC ($M = 27.82$, $ET = 11.20$) et de 18 à 51 ans ($M = 30.92$, $ET = 11.58$) dans l'échantillon EC. 45.38 % des étudiants étaient en couples (en situation de cohabitation, mariage, ou Pacs) et 54.62% ne l'étaient pas (célibataires, divorcés, ou veufs) dans l'échantillon ENC. Dans l'échantillon EC, la moitié des étudiants se répartissaient dans la situation dite de couple. La totalité des participants étudiants étaient en toute logique au-delà du niveau baccalauréat français et 53.1% des étudiants ENC n'exerçaient pas d'activité professionnelle en plus de leurs études contre 37.5% dans le groupe clinique EC.

1.5.5. Analyses statistiques

Comme nous l'avons vu, nous nous sommes inspirés de la méthodologie de Valleyrand (1989) pour construire notre protocole de collecte des données et pour sélectionner les échelles adaptées.

Dans cette section, nous analyserons les principaux résultats, selon le format plus classique qui est rencontré dans les articles soumis à des revues à comité de lecture.

Dans les études issues de la pratique clinique courante, les échantillons sont souvent réduits. Aussi, nous avons utilisé une approche complémentaire pour estimer les différents effets en utilisant soit des procédures paramétriques, soit des procédures non paramétriques et/ou des méthodes robustes en fonction des échantillons disponibles.

Les différences de scores entre les échantillons cliniques et non cliniques ont été évaluées à l'aide du test de Mann-Whitney (bilatéral à la séance 1, avec une procédure de bootstrap pour l'obtention des intervalles de confiance (IC ; Correction pour biais accélérée-Bca). Nous avons également étudié les différences dans les scores obtenus aux échelles ORS et SRS entre les conditions Feedback et non feedback.

La validité de contenu et la validité convergente ont été analysées, à la séance 1 avec les patients reçus en condition non feedback (avec les scores obtenus aux échelles ORS et SRS), en utilisant le test Rho (non paramétrique) de Spearman pour calculer les corrélations avec des intervalles de confiance obtenus par bootstrap ; et à la première occasion de mesure (test) avec le groupe d'étudiants non cliniques (ENC) pour l'échelle ORS.

La fidélité de la mesure obtenue par ces échelles a été calculée à la séance 1 avec le coefficient alpha de Cronbach, dans un groupe comprenant les patients en condition feedback et non feedback (pour les échelles ORS et SRS) et avec la population étudiante non clinique à l'occasion de mesure n°1 (test).

La validité de construit a été explorée en comparant les différences dans les scores obtenus à l'échelle ORS entre les échantillons cliniques et non cliniques à la première occasion de mesure (soit la séance 1 pour les patients ou le test pour les étudiants) et également entre la séance 1 et la séance 2 (pour les patients et les étudiants) et entre la séance 1 (entrée dans les soins) et la dernière séance disponible (pour mesurer la sensibilité au changement par le test-retest en rang de Wilcoxon), et pour les scores SRS en examinant la corrélation avec les séances suivantes. Une analyse factorielle confirmatoire (AFC) a été menée pour déterminer si la structure factorielle des versions françaises correspondait bien à la structure factorielle des versions anglaises (échelles ORS et SRS).

Des données normatives préliminaires ont été documentées (ORS, SRS, occasion 1) et les index de changement ont été calculés (pour l'échelle ORS, le seuil clinique, l'indice de changement fiable, ICF ; pour l'échelle SRS le seuil) en utilisant les formules de Truax et Jacobson (1991). Finalement, nous avons tracé en complément la courbe ROC, à partir des scores obtenus à l'échelle ORS auprès des populations cliniques et non cliniques à la séance 1 (ou l'occasion de mesure numéro 1 dite de test, pour l'échantillon étudiant non clinique).

Les taille d'effets ont été analysées à partir des recommandations de Cohen (1988 ; effet faible $d = .20$, effet moyen $d = .5$, effet important $d = .8$; effet faible $r = .1$, effet moyen $r = .3$,

effet important, $r = .5$) en considérant que les effets obtenus avec les coefficients de corrélation paramétriques et non paramétriques peuvent être analysés selon les mêmes guides d'interprétation. Nous avons utilisé pour ces analyses statistiques les logiciels SPSS (version 24), le logiciel Excel (version 16) et le logiciel Mplus (Version 7.4).

1.5.6. Validation de l'échelle Outcome Rating Scale (ORS).

1.5.6.1. Statistiques descriptives et données pré-normatives

Le Tableau 1 reproduit les indices de moyenne et d'écart type pour les scores totaux de l'échelle ORS dans chaque groupe.

Les scores moyens obtenus aux échelles ORS à la séance 1 dans les deux groupes cliniques (non feedback et feedback) sont plus bas que ceux reportés dans l'étude de validation initiale de Miller et al. (2003 ; $M = 19.6$, $ET = 8.7$) aux États-Unis, mais comparables à ceux rapportés dans les autres études de validation européenne par Janse et al. (2014 ; $M = 17.0$, $ET = 7.2$) pour l'Allemagne et par Moggia et al. (2017 ; 2018 ; $M = 18.18$, $ET = 9.91$) pour l'Espagne.

Les scores obtenus au test (première occasion de mesurage), dans l'échantillon non clinique sont plus bas que ceux rapportés par Miller et al. (2003 ; $M = 28$ $ET = 6.8$), et plus bas que ceux rapportés par Bringhurst et al. (2006 ; $M = 29.9$, $ET = 7.5$) ou ceux rapportés par Janse et al. (2014 ; $M = 29.6$, $ET = 6.0$). Les scores totaux des femmes à la séance 1 dans le groupe clinique en situation non feedback (CNF ; *Rang moyen* = 13.52) ne diffèrent pas significativement des scores obtenus par les hommes (CNF ; *Rang Moyen* = 17.43, $U = 94$, $p = .30$, $d = .21$). Les scores obtenus par les femmes dans le groupe clinique en situation feedback (CF ; *Rang moyen* = 13.06) ne diffèrent pas significativement de ceux des hommes (*Rang Moyen* = 16.42, $U = 73$, $p = .30$, $d = 0.20$). Les scores ORS obtenus à la séance 1 dans le groupe non clinique chez les femmes (ENC ; *Rang Moyen* = 67.19) ne diffèrent pas significativement de ceux des hommes (*Rang Moyen* = 56.71, $U = 960$, $p = .24$, $d = .10$).

Tableau 1. Données normatives préliminaires pour l'échelle ORS à la séance 1

	CNF		CF		ENC	
	n =28		n=28		n=130	
	(21 Femmes, 7 Hommes)		(16 Femmes, 12 Hommes)		(109 Femmes, 21 Hommes)	
	M	ET	M	ET	M	ET
Groupe	17.96	9.02	17.17	5.57	25.65	7.63
Homme	21.12	10.71	18.16	6.88	22.87	9.80
Femme	16.90	8.41	15.73	7.39	26.18	7.07

Note. CNF = groupe Clinique Non Feedback ; FC = groupe Clinique Feedback ; ENC = groupe Étudiant Non Clinique

1.5.6.2. Analyse des items et de l'échelle

Analyse de la Structure factorielle

Nous avons pu localiser une recherche en analyse exploratoire de la structure factorielle de l'échelle ORS (Smith et al., 2010 ; EFA). Les saturations pour chaque item de l'échelle ORS étaient respectivement .69, .62, .76, .98.

Nous avons testé si cette structure factorielle était répliquée dans la population française sur la population de notre étude, en réalisant une Analyse Factorielle Confirmatoire (CFA) avec les scores obtenus aux items de l'échelle, sur le plus gros échantillon disponible, l'échantillon étudiant dans sa totalité, lors de la première occasion de mesure (test ; $n = 154$). En effet, la structure factorielle est supposée être identique avec des sujets issus de la population clinique ou non clinique, seuls les niveaux de scores sur les items sont susceptibles de varier entre ces deux populations. Nous avons donc estimé un modèle en un facteur à l'aide du logiciel Mplus version 7.4 (Muthén & Muthén, 2010), en utilisant une procédure robuste (MLR) pour assurer la cohérence des analyses avec les échantillons plus réduits. Le résultat laissait apparaître une solution impropre, à savoir une matrice θ contenant une variance négative des résidus pour

l'item d ("Globalement"), ce qui est communément désigné par le terme « Heywood case », bien que ce résultat ne fût pas significatif ($p > .05$). Nous avons rencontré le même problème en réalisant une autre AFC avec les scores obtenus à l'échelle ORS sur l'Échantillon Clinique Total (ECT, composé des scores obtenus sur les items de l'ORS dans les populations en condition feedback-CF) et non feedback (CNF), et en réalisant l'analyse avec l'échantillon clinique en condition non feedback (CNF). Ce résultat ne fut par contre pas retrouvé avec l'échantillon clinique en condition feedback (CF). Ces données complémentaires seront évoquées lors de la discussion relative à cette étude.

La variance négative des résidus pour les scores obtenus à l'item d de l'échelle ORS étant petite et non significative ($-.05, p > .05$), nous avons suivi la procédure décrite par Muthen¹³ et nous avons contraint la variance de l'item d à la valeur 0. Avec cette procédure, le modèle a ajusté les données de manière satisfaisante ($\chi^2(2) = 3.89, p = .27, CFI = 1, TLI = 0.98$; $RMSEA = 0.04, 90\% IC [.00, .15]$). Les saturations pour chaque item étaient respectivement .88, .54, .60, 1. Le diagramme de l'AFC n'est pas reproduit ici au vu de la simplicité de la structure factorielle.

1.5.6.3. Analyse de la fidélité

Consistance Interne

Nous n'avons pas retrouvé de différences significatives entre les scores totaux obtenus à l'échelle ORS à la séance 1 pour les patients en condition non feedback ($n = 28, Rang\ moyen = 29.48$) et pour les patients en condition feedback ($n = 28, Rang\ moyen = 27.52, U = 364.50, p = 0.07, r = 0.06$) ; aussi nous avons agrégé ces données pour créer un échantillon clinique

¹³ <http://www.statmodel2.com/discussion/messages/9/572.html?1319030818>

total ou ECT décrit précédemment (ECT = CNF+FC, $n = 56$; Score ORS total $M = 17.57$, $ET = 8.26$).

Nous avons calculé la consistance interne à l'entrée dans les soins, dans la population clinique totale (ECT), et au test dans l'échantillon non clinique étudiant (ENC). Dans le groupe clinique (ECT) la valeur de l'alpha était très bonne au vu du nombre d'items réduits (Solution à 4 items ; $\alpha = 0.83$) et ce coefficient n'a pas été augmenté par l'utilisation d'une stratégie de pointage des items, c'est-à-dire le calcul de l'alpha avec la suppression successive de l'un des items. Dans le groupe étudiant non clinique (ENC), La solution à 4 items correspondait à une valeur identique, mais l'alpha était cette fois augmenté (.02) quand l'item a (« Bien-être individuel) était supprimé avant le calcul (solution initiale à 4 items, $\alpha = 0.83$).

Les résultats obtenus peuvent être comparés à ceux de Miller et al. (2003 ; $\alpha = .87$), Janse et al. (2014 ; $\alpha = .82$) et Moggia et al. (2017 ; $\alpha = .88$) en ce qui concerne les échantillons cliniques.

1.5.6.4. Validité de construit

Validité convergente établie pour l'échantillon Clinique (en condition non feedback CNF)

Nous avons étudié la matrice de corrélation entre les scores obtenus aux différents items de l'échelle ORS, le score total et les scores obtenus aux différentes sous-échelles du questionnaire MI-22 et son score total, en condition non feedback car il s'agissait du seul échantillon clinique de notre étude pour lequel les mesures étaient administrées concomitamment (tableau 2). Pour alléger la présentation, seules les corrélations significatives sont rapportées. Il est attendu que la corrélation soit négative pour la relation entre l'échelle ORS, où un plus haut score correspond à un plus grand degré de bien-être et pour l'échelle MI-22 où un plus haut score correspond à un plus grand degré de détresse psychologique.

Tableau 2. Corrélations obtenues entre les différents items de l'ORS et les sous-échelles de La MI-22, à la première séance de thérapie en condition non feedback. ($n = 28$)

	MI-22 DS	MI-22 RI	MI-22 RS	MI-22 Total
ORS "Individuellement" (a)	-.42[-.69,-.002]*	ns	ns	ns
ORS "Interpersonnel" (b)	ns	-.42 [-.71, -.03] *	ns	ns
ORS "Socialement"(c)	-.66 [-.86, .33]**	-.39 [-.76, .08]*	-.54 [-.81,-.17]**	-.66 [-.86, -.30]**
ORS "Globalement" (d)	-.52 [-.77, .14]**	ns	ns	-.44 [-.68, -.0.10]*
ORS "Total" (T)	-.48 [-.75, -.06]*	ns	-.38 [-.66, -.06]*	-.46 [-.72, -.10]*

Notes. MI-22 DS = Détresse subjective, RI = Relations interpersonnelles, RS =Rôle Social ; sont respectivement rapportés rs entre crochets 95% BCa IC, significativité (*.05 bilatéral ** .01 bilatéral, ns = non significatif)

Les corrélations (rho de Spearman) entre les scores à l'item a de l'ORS (Individuellement) et les scores obtenus à la sous-échelle de "détresse subjective" de la MI-22 étaient significatives ($p < .05$, effet modéré). Les corrélations entre les scores obtenus à l'item b de l'ORS (Interpersonnel) et les scores obtenus à la dimension « Relations Interpersonnelles » de la MI-22 ($p < .05$, effet modéré), la corrélation entre les scores obtenus à l'item c de l'ORS (Socialement) et la sous-échelle « Rôle Social » de l'OQ-22 étaient significatives ($p < .01$, effet important).

Les scores totaux aux questionnaires ORS (T) et MI-22 Total corrént également significativement ($p < 0.05$, effet modéré).

Au niveau des items, nous pouvons remarquer que l'item a (Individuellement) et l'item b (Interpersonnel) corrént avec chaque sous échelle respectives de la MI-22 de manière quasi exclusive, alors que pour l'item c de l'ORS (Socialement), la corrélation est significative avec chaque sous-échelles de la MI-22. Le score total à l'ORS est corrént significativement avec toutes les sous-dimensions de la MI-22 à l'exception de la sous-échelle « Relations interpersonnelles ». Nous pouvons noter que les corrélations les plus fortes sont retrouvées dans les relations entre les scores obtenus à l'item c de L'ORS (Socialement) et les scores obtenus à

la sous-échelle « Détresse subjective » de la MI-22 (effet important, $r_s = -.66$) et entre les scores obtenus à l’item c (Socialement) et le score total obtenu à la MI-22 ($r_s = -.66$).

Tableau 3. Corrélations entre les scores des items de l’ORS et les scores de l’échelle K6 dans le groupe clinique non feedback à la séance 1 (n = 28)

	K6 Total
ORS Individuellement (a)	-.55 [-.78, -.17]**
ORS Interpersonnel (b)	-.38 [-.71, .03]*
ORS Socialement (c)	-.65 [-.88, -.30]**
ORS Globalement (d)	-.65 [-.82, -.32]**
ORS Total (T)	-.63 [-.81, -.34]**

Notes. Sont respectivement rapportés r_s entre crochets 95% BCa IC, significativité (*.05 bilatéral ** .01 bilatéral, ns = non significatif

Les scores obtenus aux différents items de l’échelle ORS ainsi que les scores totaux obtenus à l’échelle K6 sont également significativement corrélés (Tableau 3). La corrélation la plus faible étant retrouvée entre les scores à item b de l’ORS (Interpersonnel) et le score total K6 ($r_s = -.38$, effet modéré) et la plus importante entre les items c (Socialement) et d (Globalement) de l’ORS et le score total à l’échelle K6 ($r_s = -.65$, effet important).

Validité convergente dans l'échantillon non Clinique ENC

Tableau 4. Corrélations entre les scores obtenus aux items de l'ORS, et les scores obtenus aux sous échelles de la MI-22 dans le groupe non clinique au test ($n = 130$)

	MI-22 DS	MI-22 RI	MI-22 RS	MI-22 Total
ORS Individuellement (a)	-.63 [-.74, -.50]**	-.52 [-.64, -.38]**	-.52 [-.64, -.36]**	-.65 [-.75, -.51]**
ORS Interpersonnel (b)	-.38 [-.53, -.20]**	-.54 [-.65, -.40]**	-.34 [-.50, -.16]**	-.46 [-.60, -.31]**
ORS Socialement (c)	-.40 [-.56, -.21]**	-.51 [-.63, -.37]**	-.47 [-.62, -.31]**	-.50 [-.64, -.34]**
ORS Globalement (d)	-.68 [-.78, -.55]**	-.59 [-.70, -.46]**	-.58 [-.69, -.44]**	-.71 [-.81, -.60]**
ORS Total (T)	-.64 [-.75, -.50]**	-.66 [-.76, -.55]**	-.59 [-.71, -.44]**	-.71 [-.80, -.60]**

Notes. MI-22 DS = Détresse subjective, RI = Relations interpersonnelles, RS = Rôle Social ; sont respectivement rapportés r entre crochets 95% BCa IC, significativité (*.05 bilatéral ** .01 bilatéral, ns = non significatif)

Tableau 5. Corrélations entre les scores obtenus aux items de l'ORS et les scores obtenus à la K6 total dans le groupe étudiant non clinique (ENC) au test ($n = 130$)

	K6 TOTAL
ORS Individuellement (a)	-.63 [-.73, -.50]**
ORS Interpersonnel (b)	-.27 [-.43, -.07]*
ORS Socialement (c)	-.38 [-.55, -.19]**
ORS Globalement (d)	-.65 [-.75, -.50]**
ORS Total (T)	-.59 [-.70, -.44]**

Notes. MI-22 DS = Détresse subjective, RI = Relations interpersonnelles, RS = Rôle Social ; sont respectivement rapportés r entre crochets 95% BCa IC, significativité (*.05 bilatéral ** .01 bilatéral, ns = non significatif)

Dans l'échantillon non clinique (Tableau 4), les corrélations (r de Pearson) entre les scores obtenus à l'item a de l'ORS (Individuellement) et les scores obtenus à la sous-échelle "Détresse subjective" de la MI-22 étaient significatives ($p < .01$, effet important), comme les corrélations entre les scores obtenus à l'item a de l'ORS (Interpersonnel) et les scores obtenus à la sous dimension « Relations Interpersonnelles » de la MI-22 ($p < .01$, effet important), ainsi

que les corrélations entre les scores obtenus à l’item c de l’ORS (Socialement) et les scores obtenus à la sous-échelle « Rôle Social » de la MI-22 ($p < .01$, effet modéré). Les scores totaux obtenus à l’échelle ORS et à l’échelle MI-22 sont également significativement corrélés ($p < 0.01$, effet important). Les scores de chaque item de l’échelle ORS sont également corrélés avec les scores obtenus à chaque sous-échelle de la MI-22. Les associations les plus importantes sont retrouvées entre les scores à l’item a de l’ORS (Individuellement) et les scores obtenus à la sous-échelle « Détresse subjective » de la MI-22, et entre l’item a de l’ORS et le score total obtenu à la MI-22 (respectivement $r = -.63$ and $r = -.64$) ; ainsi qu’entre les scores à l’item d de l’ORS (Globalement) et les scores à la sous-échelle « Détresse subjective » de la MI-22 et entre les scores obtenus à l’item d de l’ORS et les scores totaux obtenus à la MI-22 (respectivement $r = -.68$ et $r = -.71$). Avec la relation la plus forte retrouvée entre les scores totaux à l’échelle ORS et les scores totaux à la MI-22 ($r = -.71$).

Tous les scores obtenus aux items de l’échelle ORS sont également significativement corrélés avec les scores totaux obtenus à l’échelle K6 (Tableau 5), avec l’effet le plus faible et modéré pour la corrélation K6 total-ORS item b ($r = -.27$) et la plus importante pour la relation scores à l’item d de l’ORS « Globalement » et les scores K6 Total ($r = -.65$, effet important). Ces résultats sont comparables à ceux de Miller et al. (2003, 2016), Janse et al. (2014) et Moggia et al. (2017 ; 2018) dans la magnitude des effets retrouvés entre les scores des items de l’ORS, les scores globaux obtenus à cette échelle et d’autres échelles de référence. Ce résultat indique que les échelles ORS, K6 et dans une moindre mesure l’échelle MI-22, semblent évaluer des construits qui varient parallèlement et qui pourraient être reliés à une même dimension de détresse psychologique générale.

Validité critérielle

L’ORS étant destiné à mesurer un changement psychothérapeutique (état), les scores de bien-être obtenus à cette échelle devraient être prédits, dans le sens d’une relation attendue négative, par la mesure d’un trait nommé « détresse psychologique ». Aussi, nous avons testé

cette relation en procédant à une régression linéaire entre les scores ORS obtenus à l'occasion 1 (test) et les scores obtenus par l'ensemble du groupe étudiants (n=154) à la dimension Affectivité négative de l'échelle BFI. Cette relation était statistiquement significative ($\beta = -.43$, $R^2 = .19$, $p < .01$).

En outre, les scores totaux obtenus à l'échelle ORS des étudiants cliniques lors du test, (EC ; *Rang Moyen* = 63.60) ne différaient pas significativement des scores totaux obtenus à l'ORS dans le groupe étudiant non clinique (ENC ; *Rang Moyen* = 80.07, $U = 1226.50$, $p = .10$, $r = .13$). Ceci suggère que les scores ORS des étudiants qui présentent des difficultés psychologiques et qui sont en cours de traitement (psychothérapie, médication) sont statistiquement non distinguables de ceux des étudiants qui ne sont pas en cours de traitement pour un trouble psychologique. Ce résultat est attendu théoriquement puisque les individus qui reçoivent un traitement efficace sont supposés voir leur score évoluer et transiter de la région définie comme la zone clinique des scores ORS (en dessous du seuil clinique) à la région qui contient les scores non cliniques ou zone non clinique (les scores ORS au-dessus du seuil).

Sensibilité au changement

Possibilité de distinguer les individus appartenant à la population clinique et les individus appartenant à la population non clinique à partir du score total à l'ORS.

Dans la continuité des résultats de la section précédente, les scores ORS des étudiants non cliniques à la séance 1 (ENC, n = 130 ; $M = 25.65$) étaient significativement plus haut que les scores obtenus avec le groupe clinique dans sa totalité (ECT, n = 56 ; $M = 17.57$, $t(97.23) = 6.26$, $p < 0.01$, $r = .42$ ou $d = .74$).

Sensibilité au changement durant le traitement

Seuls les clients qui entamaient un premier épisode de traitement, ont été pris en compte pour cette analyse. Pour les clients en condition non feedback, le dernier score ORS total était significativement plus haut (n= 16, $M = 28.01$) que le score à l'entrée du traitement client ($M =$

18.57, $T = 123$, $p < 0.01$, $r = 0.71$). Avec les clients en condition feedback, le dernier score à l'ORS était significativement plus haut ($M = 29.15$) que le score à l'entrée du traitement ($n = 25$, $M = 17.04$, $T = 297$, $p < 0.01$, $r = 0.86$).

Ce sont des arguments en faveur de la possibilité de discriminer le statut clinique et non clinique et de suivre la progression des patients durant le traitement en utilisant le score total obtenu à l'échelle ORS.

Stabilité temporelle

Nous avons calculé la corrélation test-retest entre les occasions de mesure 1 et 2. Avec un délais comparable, le coefficient de corrélation était plus élevé pour l'échantillon étudiant non clinique (ENC, *délais moyen entre les deux mesures* = 15.92 jours, $ET = 5.31$, $r_s = .57$, $p < .01$, 95% BCA IC [.34, .74]) que pour le groupe clinique en condition feedback (CF, *délais moyen entre les deux mesures* = 13.08 jours, $ET = 7.92$, $r_s = .41$, $p < .05$, 95% BCA IC [-.05, .80]) pour lequel l'intervalle de confiance contenait le zéro. Avec un délai sensiblement plus long, la corrélation test-retest obtenue sur l'échantillon clinique en condition non feedback était non significative (CNF ; *Délais Moyen* = 26.56 jours, $ET = 14.40$, $r_s = .31$, $p > .05$, BCA IC [-.17, .66]) et l'intervalle de confiance contenait le zéro.

Le coefficient de corrélation test-retest étant donc inférieur dans les groupes cliniques, supérieur et significatif dans le groupe non clinique, les scores totaux obtenus à l'échelle ORS semblent donc varier entre deux occasions de mesure en condition de psychothérapie et rester relativement stable dans une population ne recevant pas de traitement. Ceci documente la sensibilité au changement de l'ORS pour la population clinique et l'invariance relative pour la population non clinique.

1.5.6.5. Seuil Clinique et indice de changement fiable

La méthodologie décrite par Jacobson (1991) permet de déterminer le score correspondant au seuil clinique ou *Critère c*, qui distingue les personnes appartenant à la

population clinique et celles appartenant à la population non clinique (au score seuil la probabilité d'appartenir à la population clinique vs non est de 50%). Le Critère c a été calculé, en conséquence, à la séance 1 ou au test, à partir des moyennes et des écarts types des scores ORS des étudiants « non cliniques » (ENC) et des scores obtenus par les patients constituant l'échantillon Clinique total (ECT). Le seuil clinique ainsi calculé était de 22 points (21.77), un score inférieur à celui calculé avec l'échantillon de la population américaine ($c = 25$) et inférieur au score de la population allemande ($c = 24$).

Jacobson (1991) indique également comment calculer l'indice de changement fiable (Reliable Change Index), qui correspond à la différence de scores entre deux mesures, nécessaire pour qualifier un « vrai » changement (dépassant l'erreur de la mesure et le passage du temps). Cet indice semblerait plus fiable quand le calcul s'effectue sur les données d'un échantillon non clinique, parce que la mesure des score ORS sur ce groupe n'est pas censée fluctuer de manière significative sur ce groupe entre deux occasions de mesures par simple fluctuation (Seidel et al., 2016) ; nous avons vu qu'au test-retest, la corrélation des scores sur ce groupe était importante et significative. Ce calcul donnera donc un indice plus conservateur. Nous avons utilisé la score α de fidélité de la mesure et l'écart type ET obtenu au test sur l'échantillon non clinique étudiant. L'ICF de la version française de l'ORS a été calculé à 9 points (8.68) avec ces paramètres. Il est identique à l'indice obtenu par nos collègues allemands (9 points) mais sensiblement supérieur à l'ICF de 5 points issu de la validation initiale américaine de l'échelle ORS. Afin de compléter nos données, nous avons également calculé l'ICF avec les paramètres obtenus sur l'échantillon clinique total, et trouvé un résultat assez similaire (9.39 points).

Receiver Operating Characteristic Curve (ROC)

La courbe ROC a été dessinée dans le logiciel SPSS en utilisant l'échantillon le plus important de patients à la séance 1 (ECT, $n = 56$) et l'échantillon étudiant non clinique (ENC, $n = 130$). L'aire totale sous la courbe représentait .77 et le score de 23 (22.92) nous a semblé

offrir un équilibre convenable entre la spécificité et la sensibilité (respectivement .70, .75 ; 1-sensibilité = .25) ; ce résultat correspond à une unité de plus que l'indice calculé pour le critère *c*, et à deux unités de moins que le seuil clinique américain (25 ; Miller & Duncan, 2004) et une unité de moins que le seuil clinique allemand (24 ; Janse et al., 2014) . Ces résultats vont dans le sens de ce qui est attendu d'une variation interculturelle.

1.5.7. Validation de l'Échelle d'Évaluation de la Séance-Session Rating Scale (SRS)

1.5.7.1. Statistiques descriptives -Données pré-normatives

Les scores obtenus à l'échelle SRS sont issus uniquement des échantillons cliniques où une relation thérapeutique a pu se tisser. Il s'agit des patients en condition feedback (28 patients ; score SRS *Moyen* = 34.55, *ET* = 5.17.) et non feedback (22 patients disponibles pour cette échelle ; Score SRS *Moyen* = 34.78, *ET* = 6.52) à la séance 1.

Les scores ne diffèrent pas significativement entre les conditions non feedback (CNF ; *Rang Moyen* = 23.41) et feedback (CF ; *Rang Moyen* = 28.16, *U* = 249.50, *p* = 0.25, *r* = .16), ce qui cadre avec la littérature sur l'utilisation de la SRS. En effet une étude sur l'administration de l'échelle SRS et des effets dus à la désirabilité sociale suggéraient que les scores ne différaient pas significativement si les patients s'attendaient à ce que leur feedback sur l'alliance recueilli sur cette échelle soit communiqué à leur thérapeute ou non (Reese et al., 2013). De ce fait, et afin de pouvoir maximiser les chances de détecter des effets s'ils existent, nous avons agrégé les scores obtenus par les patients dans les conditions feedback et non feedback à la première séance pour les traitements statistiques successifs. En outre, nous n'avons pas constaté de différence significative entre les scores obtenus à l'échelle SRS pour les femmes à la séance 1 (ECT ; *n* = 50 ; *Rang moyen* = 27.27) et ceux des hommes (*Rang Moyen* = 22.06, *U* = 222, *p* = .23, *d* = 0.17).

Le Tableau 6 reproduit les moyennes et les écarts types des scores obtenues à l'échelle SRS à la première séance de thérapie.

Tableau 6. Données pré-normatives obtenues à l'échelle SRS dans le groupe Clinique total (ECT) à la première séance

TCS		
<i>n</i> = 50 (17 Homme, 33 Femmes)		
	<i>M</i>	<i>ET</i>
Total	34.65	5.74
Hommes	33.25	6.87
Femmes	35.37	5.03

Notes : ECT = Échantillon Clinique Total

1.5.7.2. Analyse des items

Les recherches antérieures sur l'échelle SRS convergent vers la possibilité que l'échelle SRS possède une structure unidimensionnelle. Les items sont significativement corrélés et une analyse factorielle exploratoire (EFA) avec les scores obtenus à l'échelle SRS de Groupe (Group Session Rating Scale, GSRS) qui est basée sur le même modèle que la SRS mais est conçue pour être utilisée en thérapie de groupe, a permis de dégager une structure en un facteur (Quirk et al., 2013). Les saturations factorielles des items de la SRS a, b, c, d étaient respectivement .72, .88, .86, .90).

Pour confirmer ces résultats (AFC), nous avons estimé un modèle de mesure en un facteur avec le logiciel Mplus (version 7.4), en utilisant une procédure robuste du fait de l'échantillon réduit (MLR ; Muthén & Muthén, 2010). Les indices d'ajustement du modèle indiquent que cette solution à un facteur est acceptable ($\chi^2(2) = 0.07, p = .96, CFI = 1.00, TLI = 1.13, RMSEA = 0.00, 90\% IC .00-.00$). Les saturations pour chaque item a, b, c, d étaient respectivement (0.93, 0.87, 0.85, 0.91). Le diagramme des saturations n'est pas reproduit ici étant donné la simplicité de la structure factorielle.

1.5.7.3. Analyse de la Fidélité

La consistance interne des scores obtenus aux items de la SRS a été calculée avec l'échantillon Clinique total à la séance 1 (50 individus ayant rempli le questionnaire), et la corrélation test-retest entre les scores obtenus à la séance 1 et à la séance 2 a été également calculée ($n = 38$). Le coefficient alpha de Cronbach de la SRS était excellent ($\alpha = .93$) et n'a pas augmenté du fait de la procédure de pointage ou de test en suppression successive de l'un ou l'autre des items avant un recalcul.

1.5.7.4. Analyse de la validité du construit.

Les corrélations entre les scores obtenus aux sous-échelles de la WAI-S à la séance 1, et les scores obtenus à l'échelle SRS, en condition non feedback seulement où les échelles ont été administrées concomitamment, sont reproduites dans le tableau n°7. Nous pouvons remarquer que toutes les corrélations sont significatives ($p < .05$) avec des tailles de l'effet qui vont de $r_s = .53$ (modérément fort) à $r_s = .81$ (effet important).

Les scores totaux à l'échelle WAI-S sont également significativement corrélés avec les scores obtenus aux différents items de l'échelle SRS et son score total ($p < .01$, effet important). Ces résultats vont dans le sens que ces questionnaires évaluent un construit dont la progression est parallèle, un construit commun aux deux échelles, qui pourrait être la mesure de dimensions communes de l'alliance thérapeutique.

Tableau 7. Corrélations entre les scores à l'échelle SRS (items et score total) et les scores à l'échelle WAI-S (sous-échelles et score total) en condition non feedback à la séance 1 (n=22)

	WAI-S	WAI- S	WAI-S	WAI-S
	Relation	Buts	Taches	Total
SRS Relation	.56 [.14, .81]**	.69 [.28 .93]**	.81 [.68, .89]**	.78 [.53, .92]**
SRS Buts	.53 [.14, .80]*	.57 [.18, .87]**	.75 [.58, .85]**	.69 [.43, .85]**
SRS Approche	.59 [.14, .90]**	.70 [.36, .91]**	.82 [.63 .93]**	.78 [.51, .92]**
SRS Globalement	.61 [.18, .88]**	.72 [.39, .91]**	.77 [.53, .89]**	.77 [.51, .93]**
SRS Total	.62 [.20, .88]**	.70 [.40, .89]**	.80 [.63, .89]**	.78 [.51, .93]**

Notes. Sont respectivement rapportés, rs entre crochets 95% BCa IC, la significativité (*.05 Bilatéral, **.01 bilatéral .WAI-S = Working Alliance Inventory-Short form)

La corrélation test-retest entre les scores obtenus lors de la séance 1 et lors de la séance 2 était significative et modérée ($M = 17.23$, $SD = 10.95$, $r_s = .42$, 95% BCa, IC [.28 .93] $p < .01$). Ce résultat se rapproche de celui obtenu par Janse et al. (2014) pour les séances 1–2. De plus, théoriquement, il est attendu que l'alliance thérapeutique évolue au cours du temps tout en présentant une relative invariance de la mesure signalant une stabilité. Et ici l'effet de la corrélation est bien modéré.

Ainsi, dans le groupe de patients en condition feedback, bien que les scores SRS ne prédisent pas significativement les scores ORS des sessions suivantes ($p > .05$), nous pouvons observer une évolution positive du score moyen obtenu à l'échelle SRS de séance en séance (de la séance 1 à la séance 15) avec l'exception d'une chute de score à la séance 4. Ceci pourrait illustrer la construction d'une alliance de travail positive pendant le traitement avec le plus bas score moyen SRS mesuré à la première séance 1 (34.55) et le plus haut score mesuré aux séances 12 à 15 (> 39).

Dans le groupe de patients en condition non feedback, de manière quelque peu similaire, les scores moyens obtenus à l'échelle SRS ont évolué positivement des séances 1 à 9, à

l'exception d'une chute également à la séance 4, mais aussi à la séance 7, avec le plus bas score mesuré à la séance 1 (34,78) et le plus haut score mesuré à la séance 9 (40).

1.5.7.5. Analyse de la validité critérielle

Nous avons vu que l'alliance thérapeutique était identifiée comme un prédicteur des meilleurs résultats au traitement dans la littérature. Aussi, une séquence d'évolution positive des scores SRS a été associée à de meilleurs résultats à l'issue du traitement (Miller et al., 2004). En concordance avec les résultats évoqués au paragraphe précédent, et compte tenu de la limitation présentée par la taille de nos échantillons cliniques, nous avons exploré l'évolution des scores SRS durant les cinq premières séances de traitement en condition feedback. Cette évolution est apparue positive avec une différence de points moyenne de 1.71 unités et un écart type de 4.71, de la séance 1 à la séance 5, avec en parallèle une évolution globale positive des scores obtenus à l'échelle ORS des séances 1 à 6, et un accroissement moyen de 11.1 unités sur cet intervalle de séance ($ET = 8.79$).

1.5.7.6. Détermination d'un score seuil de l'alliance.

Nous avons collecté un total de 168 administrations de l'échelle SRS avec les patients traités dans le groupe en condition feedback, le score moyen à l'échelle SRS étant sur cet échantillon de 36.91 ($ET = 4.19$). Dans le groupe non feedback, la SRS a été administrée 65 fois, au cours de cette étude, avec un score moyen comparable de 36.07 ($ET = 4.61$).

Ceci représente un argument supplémentaire pour la possibilité d'évaluer la qualité de l'alliance que le thérapeute reçoit ou non le feedback du patient, et une première indication du score moyen attendu à l'échelle SRS en condition de traitement.

Le score au 25ème percentile dans le groupe non feedback était de 34 (33.91) et 36 (35.66) dans le groupe feedback. Ces données peuvent être rapprochées de celles obtenues par Miller & Duncan (2000, 2004) en condition feedback, avec moins de 24% des patients dans leur échantillon normatif américain qui se situaient au score seuil de 36 ou en dessous de ce

score (et moins de 9% des patients avec un score SRS de 33 points ou moins, à chaque séance). Le score obtenu sur notre échantillon en condition non feedback est proche de celui mis en évidence lors de recherches antérieures par nos collègues allemands (Janse et al., 2014), et comparable à celui obtenu par nos collègues espagnols (Moggia et al., 2017 ; 2018 ; $P(25) = 35.5$).

Afin de disposer d'un seuil un tant soit peu conservateur en regard de l'implication de l'alliance dans la qualité des traitements, et sachant que la détermination de ce seuil n'est pas exempte d'une dimension subjective, nous pourrions considérer le score seuil de 36, comme un seuil acceptable pour la population française, en lien avec la validation américaine originale de cette échelle.

1.5.8. Discussion

L'objectif principal de ce travail était de valider la version française de l'échelle ORS et de l'échelle SRS (du système PCOMS) et donc d'en étudier les qualités psychométriques en contexte de soins francophones (cliniques et non cliniques).

Les résultats sont cohérents avec nos hypothèses de départ d'équivalence globale des qualités psychométriques des deux échelles, et en cohérence également avec les autres validations étrangères du système PCOMS.

Le pré-test a confirmé la clarté des échelles et des consignes. L'utilisation aisée des mesures (de l'anglais « feasibility ») a pu être constatée puisque les échelles ont été auto-administrées aussi bien à une population de patients sans l'intervention des cliniciens, qu'à des étudiants. Nous avons également confirmé une structure unidimensionnelle aussi bien pour l'échelle ORS que l'échelle SRS.

La sensibilité au changement de l'échelle ORS, la capacité de distinguer la population clinique et non clinique, et la progression des patients pendant le traitement, comme pour la version originale a pu être constatée. La validité convergente modérée du score total à l'échelle

ORS, est convenable du fait de la brièveté de l'échelle, dans l'échantillon clinique, et plus forte dans l'échantillon non clinique va dans la direction des résultats précédents.

La validité convergente de l'échelle SRS est importante. Ce résultat est d'autant plus étonnant que l'échelle SRS est très brève.

Comme nous l'attendions, le passage d'une culture à l'autre n'a pas affecté les qualités psychométriques des échelles, mais les indices de seuils cliniques et de changement fiables, obtenus auprès des échantillons de la population française dans notre étude, ainsi que les données prénormatives diffèrent quelque peu des indices obtenus avec la population américaine pour les échelles d'origine. Ce résultat est conforme également avec les variations rencontrées au niveau de ces indices, lors de l'adaptation des mesures ORS et SRS en allemand et en espagnol. Ce qui est théoriquement consistant avec ce qui est attendu lors d'une adaptation transculturelle.

Il convient d'être précautionneux, nous le pensons, dans l'utilisation de ces résultats pour guider la pratique clinique. Il est ainsi nécessaire d'utiliser l'Indice de Changement Fiable et le seuil clinique de l'échelle ORS, en combinaison avec les données américaines, dans une perspective éclairée par un intervalle de confiance, en ce qui concerne notamment le seuil qualifiant le passage du statut clinique à non clinique (Cut-off), et le score nécessaire pour qualifier de fiable un changement constaté entre deux occasions de mesure sur l'échelle ORS et associé à un bon résultat thérapeutique. En effet, les scores qui se situent à proximité du seuil sont statistiquement moins discriminables que les scores qui en sont éloignés, puisqu'au seuil clinique, la probabilité d'appartenir à la population clinique ou à la population témoin, est équivalente.

Il est toujours possible de compléter l'évaluation offerte par le système PCOMS en complétant l'évaluation, notamment en début et en fin de traitement par des mesures plus longues et plus sensibles au changement comme le questionnaire OQ-22 ou le questionnaire

OQ-45 et qui comprennent également des items pour repérer les conduites à risque et les problématiques d'abus de substance.

Toutefois, il est nécessaire de le noter, les recherches documentent que le niveau de détresse psychologique à l'entrée dans les soins est un prédicteur fiable des résultats finaux et que les clients qui présentent un plus haut niveau de détresse psychologique au début du traitement sont plus à même de retirer des bénéfices réels des soins (Bertolino, Bargmann & Miller, 2012).

Nous l'avons vu, les scores recueillis à l'échelle ORS sont bien corrélés avec d'autres échelles mesurant la détresse psychologique et peuvent être directement utilisés pour déterminer ce degré de détresse.

Limitations de cette recherche.

Cette étude présente bien évidemment des limitations : les échantillons constitués sont petits, bien que cela soit classique dans les études impliquant la participation de patients issus de la pratique clinique de terrain et de ce fait nous avons utilisé des tests non paramétriques à chaque fois que cela était nécessaire, ce qui peut amener à des résultats plus conservateurs et à des tailles de l'effet réduites. Il faut tout de même noter que les standards suggérés par Cohen en termes de taille des effets, procèdent en partie d'une démarche empirique, que l'auteur a édicté des mises en garde quant à leur adoption rigide, et que ces seuils peuvent se révéler particulièrement restrictifs au vu des tailles de l'effet rencontrés dans les recherches en sciences humaines. Pour Gignac & Szodorai (2016), il s'agirait en ce sens de considérer qu'une taille de l'effet de .30 correspondrait plus à des effets retrouvés dans la perspective des différences interindividuelles et correspondrait en conséquence à un effet relativement large.

En ce qui concerne l'échelle ORS, la présence d'une variance de résidus négatifs dans la matrice thêta produite lors de l'AFC, peut nous alerter sur ce qui est communément appelé un « Heywood Case ». Une des explications logiques à ce résultat peut tenir du phénomène de

multicolinéarité. En effet, dans la construction de l'échelle ORS, l'item d (Globalement) est une sorte de mesure composite recueillie après la cotation des items a, b, c, qui permettrait de dégager une tendance « globale » de bien-être, dans l'expérience du patient. En ce sens la multicolinéarité entre ces items et l'item d est possible. Les items corréleraient d'ailleurs significativement entre eux puisque la consistance interne est bonne en adoptant un format d'analyse unidimensionnelle. Cette multicolinéarité pourrait affecter la représentativité de ce modèle en un facteur. En outre, il nous paraît utile de rappeler que l'échelle ORS a été administrée de manière quelque peu différente dans nos échantillons cliniques et non cliniques issus de l'étude de validation initiale des échelles (en condition non feedback) et dans la population issue de notre pratique libérale de l'étude complémentaire (en condition feedback). En effet, afin de standardiser au maximum l'administration des questionnaires, et de répondre aux exigences de la commission internationale des tests, la solution informatique fut retenue puisque le but n'était pas de tester les effets du feedback, mais de déterminer les qualités de la mesure véhiculée par les échelles, et de laisser les patients s'autoévaluer et évaluer l'alliance thérapeutique en toute autonomie ; c'est à dire sans influence de leur thérapeute (sauf en cas de bug manifeste de la plateforme où le thérapeute pouvait intervenir et donner de l'aide). En condition feedback, les patients ont donc été guidés dans leur utilisation des questionnaires, afin de pouvoir créer une bonne « culture de feedback » susceptible de valoriser leur parole et de permettre de recueillir plus de feedback (néгатif) utile. Ceci correspondant à la pratique « bonafide » de l'utilisation des échelles ORS et SRS, qui s'inscrit dans une pratique clinique, où le feedback permet d'ajuster le traitement, la pratique FIT. Dans cette condition, nous étions en effet plus intéressés par l'influence du feedback sur le traitement que par la validité de la mesure. En condition non feedback, par contre, les participants n'ont pas eu d'autre guidance que les instructions de l'échelle ORS et du fait de la similarité entre l'item d de l'ORS (Globalement) et l'item a de l'ORS (Individuellement), les participants ont pu avoir plus de difficultés à interpréter les différences entre ces deux items, et à y répondre. Nous avons

constaté ce phénomène, au long cours, sur plus de dix ans d'utilisation clinique des questionnaires ORS et SRS (dans la démarche FIT) avec notre patientèle. Nous pouvons également ajouter qu'au pré-test, bien que l'item d ait été noté comme suffisamment clair (avec un score moyen supérieur à 4 sur 7), il s'agissait de l'item recevant la note moyenne la plus basse. En condition Feedback, avec la guidance du thérapeute, les participants ont peut-être pu plus facilement intégrer le fait que l'item d (Globalement) correspondait à une évaluation faite après avoir répondu aux questions a, b, c.

Nos données semblent aller dans cette direction. La corrélation entre les scores de l'item ORS a (Individuellement) et ORS d (Globalement) est plus importante en condition non feedback dans le groupe clinique ($rs = .90$ [.73, .97], $p < .01$) et en condition feedback dans le groupe d'étudiants non cliniques ($rs = .86$ [.81, .90], $p < .01$), et plus faible dans le groupe clinique issu de la pratique privée en condition feedback ($rs = .70$ [.34, .92], $p < .01$). En condition feedback les scores aux items a et d semblent donc mieux différenciés. Ceci laisse à penser, en conjonction avec des recherches précédentes que l'échelle ORS pourrait éventuellement être présentée en utilisant seulement le score total sommé des trois premiers items (Campbell et al., 2009) ; mais dans ce cas, le risque d'un modèle juste identifié-saturé va se poser quant à la structure factorielle.

1.5.9. Conclusion

La pratique du Routine Outcome Monitoring est nous l'avons vu susceptible d'améliorer la qualité et l'efficacité des soins psychothérapeutique et correspond à une pratique dite « responsable » (accountability) de la thérapie individuelle et de la thérapie de couple (Miller et al., 2006 ; Anker, Duncan & Sparks, 2009).

Au regard des qualités psychométriques des échelles ORS et SRS du système PCOMS adaptées en français, l'utilisation clinique de ces échelles devrait, selon nous être favorisée. Le processus dénommé FIT pourrait favoriser l'allègement de la charge des cliniciens tout en les

aidant à prévenir les ruptures de l'alliance et les abandons de soins dans la pratique dite de terrain (Bickman, Kelley, & Athay, 2012).

Par extension, cette méthodologie, pourrait amener à des éléments concrets de réponses aux enjeux courants concernant la régulation et la délivrance des psychothérapies dans les pays francophones, car il est possible d'utiliser ces outils pour suivre l'efficacité des soins et la documenter. Plus spécifiquement cette méthodologie pourrait permettre (Cazauvieilh, 2013, 2014, 2015) : 1) Le suivi des résultats dans les conditions cliniques de la pratique de terrain, afin d'ajuster le traitement dans un continuum d'interventions de santé, ce qui correspond à une opérationnalisation de la pratique dite « responsable » et à une mise en place des recommandations sur les pratiques basées sur les preuves de l'APA (APA, Task force, 2006). 2) De documenter l'efficacité de la pratique de terrain en referment un fossé entre la recherche et la pratique en psychothérapie, pour faciliter la démonstration de l'efficacité des cliniciens et des institutions de soins auprès des assurances de santé qui sont susceptibles de couvrir les frais des séances et 3) De permettre aux cliniciens et aux institutions de travailler au développement de leur efficacité en adoptant une pratique délibérée (micro et macro) basée sur le ROM.

Ces trois points sont d'autant plus importants que des expérimentations actuelles visant au remboursement des psychothérapies sont en train d'être menées en France. A ce sujet il est nécessaire de rappeler qu'au-delà des preuves de nature économique issues des calculs relatifs au rapport coût-efficacité bénéfique de la psychothérapie illustrée par des travaux récents (Dezetter et al., 2014), la psychothérapie est une approche (très) efficace pour des problématiques variées, sauf quand elle ne l'est pas ! Et qu'elle nécessite, dans ces conditions, des adaptations et des corrections, ce que vise la pratique ROM. Ceci implique donc l'utilisation d'outils de ROM tels que le système PCOMS.

1.6. Étude pilote des relations entre la capacité à décoder les signaux non-verbaux et les performances des psychothérapeutes, réalisée auprès d'un groupe de thérapeutes anglo-saxons enregistrant leurs résultats dans le cadre du Routine Outcome Monitoring : l'étude Non Verbal - Outcome (NVB-O study)

1.6.1. Origines de la démarche

1.6.1.1. Synthèse de la littérature

Nous avons évoqué à plusieurs reprises dans ce travail de thèse, qu'en dépit de l'efficacité générale de la psychothérapie (Lambert & Ogles, 2004), la différence d'efficacité constatée dans la pratique clinique semblait plutôt reliée aux thérapeutes, plutôt qu'à l'efficacité différentielle entre différentes approches *bonafide* qui sont reconnues comme efficace pour une problématique donnée. Cette différence dans l'efficacité des thérapeutes représenterait une part de variance très importante dans les résultats et constituerait la portion principale de la variance de l'efficacité des soins en comparaison avec la portion de la variance reliée au traitement spécifique délivré (Wampold, 2001 ; Wampold & Imel, 2015). Ainsi, la part de la variance due au thérapeute pourrait être équivalente à six fois l'importance de la part de la variance liée à la technique spécifique (Bertolino, Bargmann & Miller, 2012).

Nous avons également vu dans les recherches que la capacité des thérapeutes à établir une relation de collaboration satisfaisante avec leurs patients était un des prédicteurs le plus important de l'efficacité des soins (Baldwin, Wampold & Imel, 2007), avec une alliance thérapeutique évaluée côté patient qui serait plus prédictive des résultats psychothérapeutiques (Bachelor & Horvath, 1999 ; Horvath, & Bedi, 2002). De manière encore plus spécifique, et comme nous l'avons illustré dans notre proposition d'un modèle générique des effets dus au feedback en psychothérapie, l'alliance thérapeutique prédirait l'engagement des clients dans les soins, et l'engagement des clients dans les soins serait associé aux meilleurs résultats. (Orlinsky & al., 2004).

1.6.1.2. Influence des capacités relationnelles des thérapeutes

Des données issues de recherches préliminaires apportent des arguments encore plus spécifiques à ce propos en soulignant que les capacités du thérapeute à faciliter les processus interpersonnels dans la psychothérapie seraient reliés aux résultats. En effet, Anderson, Ogles, Patterson, Lambert., & Vermeersch (2009) ont montré dans une étude utilisant un échantillon de données archivées, issue de la pratique d'un centre de counseling universitaire, que les résultats étaient reliés à la capacité des thérapeutes à faciliter le processus thérapeutique par un ensemble d'habiletés relationnelles, cet ensemble de capacités étant désigné sous le terme d'habiletés des thérapeutes à faciliter les processus interpersonnels (ou Facilitative Interpersonal Skills - FIS). Les FIS avaient été évaluées dans cette étude à l'aide d'une tâche spécifique pour mesurer la performance effective et non une performance évaluée par simple mesure auto-rapportée.

Ainsi, dans un échantillon composé de 1141 patients (*âge moyen* = 23.0 ans ; *ET* = 4.1) suivis par 25 thérapeutes et ayant reçu au moins trois séances de thérapie, la progression des patients en cours de traitement était évaluée par le questionnaire OQ-45 qui mesure, comme nous l'avons déjà évoqué, la détresse psychologique relative à trois domaines de fonctionnement (Détresse subjective, Interpersonnelle, et Sociale), même si la structure a une dimension et le score global obtenu à ce questionnaire est plus communément utilisé dans la pratique. Les habiletés relationnelles des thérapeutes étaient aussi bien évaluées, dans cette recherche avec une mesure auto-rapportée, la Social Skills Inventory (SSI, Riggio, 1986) qu'à l'aide d'une tâche de performance mesurant les FIS. Dans cette tâche, il s'agissait pour les thérapeutes de répondre, comme s'ils étaient eux-mêmes les thérapeutes, à des questions concernant 8 segments de vidéos reproduisant des situations prototypiques problématiques rencontrées en psychothérapie, et qui correspondaient à des problématiques interpersonnelles. Les thérapeutes étaient ensuite évalués sur cette tâche de performance concernant les habiletés à faciliter les relations interpersonnelles (FIS) sur dix indicateurs : la fluence verbale, la capacité

de persuasion, l'expression émotionnelle, les habiletés quant à l'alliance, l'empathie, la chaleur (humaine), la focalisation sur les problèmes et la capacité à susciter de l'espoir, avec un format d'évaluation suivant une échelle de Likert en 5 points où les plus hauts scores étaient associés aux meilleures habiletés. Bien que le design de cette étude était rétrospectif, sachant que les données relatives aux patients avaient été recueillies avant d'évaluer les thérapeutes par la tâche de performance, du fait de l'utilisation d'une base de données archivées, les habiletés à faciliter le processus interpersonnel de la thérapie sont considérées comme un construit stable, en l'absence d'entraînement formel, et ce design a ainsi permis d'utiliser une analyse en modèle multiniveau (Multilevel Modeling - MLM ou Hierarchical Level Modeling- HLM, nous détaillerons plus avant ce modèle dans la section statistique de cette étude) pour évaluer si la performance liée à la tâche FIS expliquait une portion significative de la variance des résultats dus au traitement. Cependant la séquence temporelle du protocole n'autorisait pas à proprement parler à extrapoler une quelconque séquence causale puisqu'ici l'effet précédait la cause.

Les relations entre les capacités FIS et les résultats ont été explorées dans une autre étude, cette fois ci contrôlée et randomisée et plus récente, où il s'agissait d'évaluer l'efficacité de thérapeutes qui avaient été sélectionnés quant à leurs capacités FIS et à leur niveau de formation (Anderson et al., 2015). Les auteurs précisent que pour les besoins de cette étude, les thérapeutes et les patients n'étaient pas totalement conformes à la définition classique de ces termes : la population de patients n'était pas constituée par des individus qui sollicitaient activement une aide psychothérapeutique, mais était constituée d'étudiants recrutés sur la base de la sévérité de leur évaluation psychologique ; certains thérapeutes « sélectionnés » n'avaient en outre pas reçus de formation clinique formelle, puisqu'ils s'agissait de doctorants de disciplines différentes (par exemple la psychologie expérimentale, l'histoire...). Malgré ces limitations, les données de cette étude suggèrent que les thérapeutes qui étaient notés comme les plus compétents en termes d'habiletés FIS obtenaient de meilleurs résultats (des changements plus importants de séance en séance), et de meilleurs scores d'alliance thérapeutique. Il est à noter

que dans cette étude les différences dans la formation n'ont pas été associées à des différences dans les résultats ou dans l'alliance thérapeutique.

1.6.1.3. Limitation des démarches de recherches antérieures

Nous pouvons remarquer que les habiletés FIS, telles qu'elles ont été définies dans ces études, sont principalement liées aux comportements verbaux des thérapeutes (notamment la fluence verbale, la capacité de persuader, et la capacité à instiller de l'espoir) et ne sont donc pas spécifiquement reliées à des compétences non verbales ou à la capacité des thérapeutes à décoder des signaux non verbaux, habiletés qui pourraient être elles aussi mesurées à l'aide de tâches de performance.

Pourtant, nous l'avons vu, les comportements non-verbaux des patients et des thérapeutes sont de première importance. Une portion significative des messages transmis durant la communication sont de nature non-verbale ou para-verbale. La psychothérapie est principalement la forme spécifique d'une interaction ou d'une relation d'influence entre un thérapeute et son patient, où le thérapeute a pour rôle de stimuler les capacités innées de guérison du patient (Frank & Frank, 1991) ; cette relation opère donc à partir d'une forme spécifique de communication, qui implique l'échange d'informations verbales et non-verbales.

Pour mémoire, bien que la communication verbale soit le premier focus des recherches en psychothérapie, la communication non-verbale représente une source majeure d'information, qui selon les estimations varierait avec de 65% à 95 % des signaux qui pourraient ne pas être transmis par les mots (Friedman, 1978). En effet, la "dimension cachée " de l'information (Hall, 1969), la portion non-verbale de la communication comprend toutes les informations qui ne sont pas transmises par les mots ou les symboles écrits (Knapp, 1972), ou qui transite par des médias qui excluent les mots (Mastumoto et al., 2013). Cette dimension non verbale peut véhiculer des éléments d'information très significatifs à propos de l'état du patient, qui peuvent compléter l'information verbale recueillie par les thérapeutes (Ekman,

O'Sullivan & Frank, 1999).

De plus, la communication non-verbale étant « incarnée » (« embodied »), chaque partie du corps et du fonctionnement corporel peut potentiellement agir comme un vecteur d'informations (par exemple, la position du corps, la gestuelle). Ceci implique que des informations essentielles puissent transiter par les canaux para-verbaux, le ton de la voix et son rythme ; les expressions faciales par exemple représentent une part très importante de la communication. Le visage est en effet particulièrement reconnu comme une source principale d'informations, en véhiculant des expressions faciales automatiques et par conséquent, moins filtrés par les individus, de manière instantanée, et qui peuvent amener à une réduction de la complexité des interactions sociales (Matsumoto & Hwang, 2013). Pour schématiser, nous pourrions dire que les individus qui peuvent se comprendre, par exemple en décodant les signaux non verbaux, peuvent mieux interagir.

En tant qu'êtres humains, nous l'avons vu, les psychothérapeutes sont naturellement « pré-câblés » à décoder les sept émotions primaires qui s'affichent sur le visage par le biais des expressions faciales (la joie, la tristesse, la peur, la colère, la surprise, le dégoût, et le mépris). Ces expressions semblent être universelles même si elles sont modulées par le contexte, notamment le contexte social. Nous avons également vu que ces émotions prototypiques connues comme des macro-expressions, s'affichent sur toute la surface du visage avec une intensité et une durée supérieure aux expressions partielles, micro ou subtiles, qui sont affichées avec une moindre intensité et sur des régions plus discrètes du visage (Matsumoto & Hwang, 2013). Dans les contextes sociaux, comme dans le contexte de l'interaction psychothérapeutique, les individus ont tendance à réprimer ou à masquer leurs réactions négatives (Hill et al. 1992), et c'est justement dans ces contextes où les individus ressentent le besoin de cacher leurs émotions, que les micro-expressions sont supposées se manifester (Ekman, 2003 ; Matsumoto & Hwang, 2013 ; Porter & Ten Brinke, 2008).

Nous avons également évoqué le fait que ces micro-expressions pourraient être reliées

à des contenus négatifs que les patients essaieraient de camoufler pendant les séances, et que ces micro-expressions pourraient agir comme des signaux naturalistes de feedback, signaux qui pourraient être perçus et décodés par les meilleurs thérapeutes. Ekman, O'Sullivan & Frank ont ainsi démontré que les individus différaient quant à leurs capacités à détecter les expressions émotionnelles micro et subtiles ; et que certains groupes d'individus, plus spécifiquement les individus intéressés par la détection du mensonge pourraient être plus efficace dans la détection de ces signaux. Un lien a été établi entre la capacité de décoder les émotions et la capacité à détecter le mensonge (Warren, Schertler & Bull, 2009). La capacité à détecter les émotions est notamment reliée à la capacité empathique, et l'empathie est considérée comme un des facteurs communs de la psychothérapie qui expliquerait le plus de la variance des résultats dus au traitement (Wampold & Imel, 2015).

Les individus diffèrent donc dans leur capacité à discerner et à décoder les messages non verbaux de la communication (Roter et al., 2006). Ceci nous a amené à formuler l'hypothèse que la capacité des psychothérapeutes à discerner et à décoder les informations non verbales, puissent être reliées à une portion significative de la variance des résultats dus au traitement, phénomène qui selon nous a été jusqu'à présent, insuffisamment étudié.

Pour soutenir encore plus avant cette hypothèse, nous pouvons citer les travaux d'Anderson et al. (2012), qui ont étudié l'aspect dysfonctionnel de l'hostilité dans les relations interpersonnelles entre un thérapeute et son patient, dans un contexte de soins psychodynamiques. Les auteurs suggèrent que ces comportements puissent impliquer aussi bien des expressions verbales que non verbales, et que ces comportements pourraient impacter significativement le processus psychothérapeutique. Nous pouvons citer également les travaux de Hill et al. (2016) sur l'entraînement des étudiants à développer des capacités aidantes, qui suggèrent d'évaluer les habiletés interpersonnelles en utilisant des protocoles composites, avec à la fois des mesures auto-rapportées, le détail des expériences aidantes, l'utilisation de mesures basées sur la performance, et de mesures non verbales ; ces auteurs mettent l'accent sur le fait

que les mesures des capacités interpersonnelles auto-rapportées ne sont pas suffisamment précises et sont donc insuffisantes pour évaluer les capacités interpersonnelles. Nous pensons dans cette optique, que les mesures auto-rapportées des compétences sociales ou relatives aux habiletés non-verbales ne peuvent pas éclairer de manière satisfaisantes les relations entre ces habiletés et les résultats des traitement psychothérapeutiques en condition de feedback naturaliste ou instrumental. Et qu'il est donc nécessaire, soit de disposer de tâches de laboratoires, soit de questionnaires reposant sur des tâches de performances validés à cet effet.

Les messages non verbaux et le feedback issu du Suivi Continu des Résultats au traitement (ROM)

Comme la variabilité des performances des thérapeutes est substantielle, et que les thérapeutes ont des difficultés à détecter et à prédire la détérioration probable de leurs patients et les abandons pendant le traitement, la méthodologie du Traitement Orienté par le Feedback a été élaborée, nous l'avons également vu, pour aider le praticien à adapter le traitement en temps réel. Pour rappel, selon Worthen & Lambert (2007), le feedback recueilli auprès des patients pourrait impacter les résultats en apportant aux cliniciens des informations qu'ils avaient originellement ignorées ou auxquelles ils n'auraient pas accordé une juste importance. Ce qui nous a laissé supposer que ces informations puissent être reliées à des signaux non verbaux, pour lesquels les thérapeutes pourraient avoir des capacités de discrimination variables.

Aussi, la pratique ROM semble principalement efficace pour les patients qui initialement ne tiraient pas de bénéfices réels de leurs traitements (patients qui n'étaient pas « en bonne voie de guérison » pour NOT ON TRACK – ou NOT) ce qui suggère que le signal véhiculé par le feedback instrumental de la pratique ROM puisse agir comme un signal négatif au potentiel correctif. Les thérapeutes ne semblent par ailleurs pas apprendre de la réception de ce signal de feedback, le feedback semblant également bénéficier aux psychothérapeutes originellement « les moins efficaces », pour lesquels un accroissement substantiel de

l'efficacité pourrait être obtenu (atteignant l'efficacité de thérapeutes naturellement plus talentueux), sans autre entraînement plus formel à la psychothérapie, que de bénéficier de la possibilité de discriminer et d'utiliser ce signal négatif révélateur d'information importante. Ceci nous laissant supposer que le feedback instrumental puisse agir comme un complément ou un remplaçant du feedback naturaliste, pour lequel les thérapeutes pourraient avoir des capacités de détection différentes ; c'est-à-dire, avoir accès à une dimension cachée de l'information, qui est pourtant présente dans les séances de psychothérapie, ce qui va dans le sens du modèle générique du feedback en psychothérapie que nous avons suggéré.

Toutes ces données nous amènent à supposer qu'avoir accès à des éléments cachés d'informations ou à des éléments difficilement discriminables, pourrait permettre aux thérapeutes de changer l'orientation du processus de soins, en temps réel. Le feedback instrumental agirait ainsi comme une lentille grossissante, maximisant le potentiel de réception et d'utilisation de ces informations par les thérapeutes ; ces signaux étant naturellement présents dans les interactions psychothérapeutiques mais difficilement perceptibles pour tout un chacun. Comme l'alliance prédit les résultats, le feedback instrumental négatif pourrait exercer un effet de tampon entre la relation d'influence négative de la communication non verbale négative, l'alliance, et les résultats. Dans ce sens le feedback instrumental modérerait la relation négative entre la communication non verbale négative, l'alliance, et les résultats.

Dans cette optique, cette étude vise à soumettre à l'épreuve des faits nos hypothèses en analysant les résultats obtenus par des psychothérapeutes en condition de feedback instrumental et leurs habiletés à décoder la communication non-verbale.

1.6.1.4. Hypothèses relatives aux habiletés à percevoir les signaux non verbaux en psychothérapie.

Ces données nous ont amenés à postuler plusieurs hypothèses quant à la relation entre les habiletés des cliniciens à percevoir une dimension cachée (non-verbale) de la communication, les résultats de la psychothérapie en général, et les effets des pratiques

Feedback Informed Treatment en particulier.

Pour cette étude nous avons posé l'hypothèse principale que les habiletés des thérapeutes à décoder les signaux non verbaux sont significativement reliées aux résultats de la psychothérapie (H1).

Nous émettons l'hypothèse plus spécifique que la capacité à décoder différents signaux non verbaux (macro expressions émotionnelles et micro expressions émotionnelles) et par le biais de différents canaux (corps, visage, voix) puisse être significativement associée aux résultats de la psychothérapie (H2).

Nous postulons également que la capacité à détecter les micro expressions est plus significativement associée aux meilleurs résultats que la capacité à détecter les macro-expressions (H3)

Nous émettrons aussi l'hypothèse que la capacité à détecter les micro-expressions négatives est plus spécifiquement liée aux meilleurs résultats que la capacité à détecter les micro-expressions positives ou les micro expressions en général (H4).

Enfin, nous supposons qu'en condition de feedback instrumental, la capacité à détecter les signaux non verbaux est associée à de meilleurs résultats (H5). Cette hypothèse va dans le sens d'un effet correcteur ou supplétif du feedback (effet modérateur). L'hypothèse alternative qu'en condition de feedback instrumental il n'existerait plus d'association significative entre la capacité à détecter les signaux non verbaux et l'efficacité due au traitement, ce qui correspondrait plutôt à l'hypothèse de l'effet palliatif ou de remplacement du feedback naturaliste par le feedback instrumental, et donc à l'hypothèse de l'hypothèse médiatrice des effets dus au feedback instrumental.

1.6.2. Méthodologie générale de notre étude

Afin d'étudier le rôle de la reconnaissance des signaux non verbaux dans les résultats dus au traitement, nous avons cherché les possibilités d'utiliser les données récoltées par des psychothérapeutes enregistrant leurs résultats dans leur pratique clinique en condition de ROM.

A cet effet, nous avons reçu l'autorisation de Bill Andrews de présenter notre recherche et d'inviter les thérapeutes de son réseau de recherche et de pratique (Practice Research Network, Royaume Uni - Pragmatic Research Network) à y prendre part. Les thérapeutes de son réseau enregistrent en effet leurs résultats en continu à l'aide d'un logiciel disponible sur internet nommé « Pragmatic Tracker », et peuvent ainsi afficher les données récoltées pour adapter leur traitement en fonction de la réponse des patients. Le feedback est obtenu à partir des données brutes de résultats aux questionnaires, qui peuvent être individualisés d'un thérapeute à un autre et d'un patient à un autre, et il n'y a pas d'information complémentaire fournie par un algorithme sur la détérioration des patients pendant les soins ou d'indication de conseil sur la marche du traitement. En ce sens, les thérapeutes reçoivent un feedback minimal comparable à celui d'un clinicien qui enregistrerait ses résultats avec un questionnaire papier, en continu, tout en bénéficiant d'une administration facilitée, et de la possibilité d'exploiter ses données à tout moment.

Pour inciter les thérapeutes à participer à notre recherche, nous avons annoncé aux thérapeutes, que s'ils prenaient part à notre étude, ils recevraient, à la fin de la recherche, une formation en ligne sur les liens entre la communication non verbale et l'efficacité des psychothérapies. Douze thérapeutes ont initialement donné leur accord pour participer à cette recherche, et onze thérapeutes ont finalement complété le protocole total.

1.6.2.1. Critères d'inclusion

Les données archivées des patients de ces onze thérapeutes, sur la première et la dernière séance de thérapie ont été récupérées auprès de Bill Andrews. Les patients ayant initialement

donné leur accord à leurs cliniciens, pour l'enregistrement des données avec le logiciel « Pragmatic Tracker » et pour l'utilisation des données archivées, il ne fut pas nécessaire de les contacter pour leur adresser un formulaire de consentement éclairé. Toutefois, nous avons enregistré notre collecte et notre traitement des données auprès de la CNIL. L'inclusion s'est donc faite sur la base des données présentes dans le fichier des thérapeutes au moment de la participation à notre étude. Ces thérapeutes ont signé un consentement éclairé leur expliquant que le protocole explorait leur capacité de reconnaissance des émotions au travers de différentes tâches, et qu'il leur serait également posées des questions sur leur statut professionnel, leur pratique, et sur leur façon de répondre à leurs émotions dans leur travail. Le consentement éclairé leur expliquait le protocole et la garantie de la protection de leurs données.

1.6.2.2. Procédures

Chaque thérapeute a pris part à cette étude par le biais d'un questionnaire dispensé à l'aide de l'application de questionnaires en ligne Limesurvey (Carsten Schmitz ; 2003) avec un accès sécurisé par code individuel. Ayant changé d'affectation de laboratoire de recherche, au cours de notre thèse, il ne nous fut pas possible d'utiliser la plateforme Limesurvey de l'université qui avait été utilisée pour l'étude de validation des échelles PCOMS. Aussi, nous avons assuré la mise en ligne et le déploiement de l'application Limesurvey sur un site internet professionnel que nous avons créé et dédié à nos recherches.

Dans le processus de recrutement des thérapeutes, nous n'avons pas eu accès à leur identité ou à des informations personnelles de nature à les identifier, et la gestion des participants à l'étude a été entièrement menée par Bill Andrew, à qui les codes de participation et d'accès sécurisé à la plateforme Limesurvey ont été confiés. Bill Andrews n'a pas eu accès aux réponses des participants à notre enquête.

Les thérapeutes ont rempli un formulaire de consentement éclairé qui est reproduit en annexe de cette thèse (Annexe 12) dans le but de participer à l'étude et de répondre à une série

de questions et de tâches pour une durée approximative de 45 minutes. L'Étude comprenait également des questions spécifiques de contrôle (données socio-démographiques) et des questions sur le statut professionnel des thérapeutes.

Une partie de l'étude a été administrée à partir du site spécifique de l'IRPPSY dédié à l'exécution de la plateforme Limesurvey à des fins de recherches, et les thérapeutes ont été redirigés depuis ce site vers une seconde partie de l'enquête qui comprenait deux autres questionnaires administrés depuis un consortium universitaire suisse (Swiss Center For Affective Science). La gestion de l'accès par code a permis d'anonymiser les données récoltées. Les deux questionnaires utilisés dans la seconde partie de l'étude impliquaient que les données brutes des thérapeutes puissent être analysées par le consortium universitaire suisse à des fins de recherche, ce qui fut mentionné dans le formulaire de consentement éclairé.

1.6.3. Échelles administrées et données recueillies

1.6.3.1. Le Questionnaire Clinical Outcomes in Routine Evaluations-10

Ce questionnaire (CORE-10 ; Barkham et al. (2013)), est une échelle auto-administrée composée de 10 questions permettant d'évaluer trois domaines de fonctionnement (le fonctionnement général, le fonctionnement social et les relations proches) et différentes problématiques présentées par les patients (la dépression, l'anxiété, les traumas, les problèmes physiques, le risque pour soi-même) au cours de leur psychothérapie. Chaque item reçoit une réponse sur une échelle de Likert en 5 points, avec des scores de 0 à 4. Le score total à l'échelle est obtenu en additionnant les scores des items pour un score total compris entre 0 et 40. De plus hauts scores à l'item ou à l'échelle indiquent de plus hauts niveaux de détresse psychologique. Dans cette étude, l'échelle CORE-10 a été utilisée pour évaluer les résultats psychothérapeutiques et administrée par certains thérapeutes à l'aide du logiciel Pragmatic Tracker.

1.6.3.2. Le questionnaire Patient Health Questionnaire-9

Ce questionnaire (PHQ-9 ; Kroenke, Spitzer & Williams, 2001 ; Spitzer, Williams & Kroenke, 2000) est une échelle de mesure auto-administrée composée de 9 items qui visent à évaluer les critères de la dépression rapportés par le DSM-IV. Chaque item reçoit une réponse sur une échelle de Likert en 4 points, avec des scores compris entre 0 à 3. Le score total à l'échelle PHQ-9 est obtenu par l'addition des scores individuels pour un score total compris entre 0 et 27. De plus haut scores indiquent de plus hauts niveaux de dépression. L'échelle PHQ-9 est utilisée dans cette étude comme un questionnaire visant à analyser les résultats psychothérapeutiques, et a été administrée par certains thérapeutes de notre étude pour évaluer les résultats psychothérapeutiques à l'aide du logiciel Pragmatic Tracker.

1.6.3.3. Le questionnaire Generalized Anxiety Disorders -7

Ce questionnaire (GAD-7 ; Spitzer, Kroenke K, Williams & Löwe B., 2006), est une échelle auto-administrée, composée de 7 items visant à évaluer la sévérité des symptômes d'anxiété (généralisée). Chaque item reçoit une réponse sur une échelle de Likert en 4 points avec des scores compris entre 0 et 3. Le score total étant obtenu par l'addition des scores au différents items est compris entre 0 et 21. De plus hauts scores indiquent des niveaux d'anxiété plus élevés. L'échelle GAD-7 a été utilisée dans cette étude pour évaluer les résultats et administrée par certains thérapeutes à l'aide du logiciel Pragmatic Tracker.

1.6.3.4. Le questionnaire Work-Related Acceptance and Action Questionnaire

Ce questionnaire (WAAQ, Bond, Lloyd, & Guenole. 2013), est un questionnaire auto-administré, composé de 7 items et destiné à évaluer la flexibilité psychologique dans le contexte du travail. La flexibilité psychologique étant définie comme la capacité à persister ou à changer son comportement, notamment en présence de ressentis et pensées difficiles, dans le but de réaliser des actions qui servent les valeurs (Hayes et al., 2012). Chaque item reçoit une réponse sur une échelle de Likert en 8 points avec des scores de 0 à 7. Le score total à la WAAQ est donc compris entre 0 et 49. Les plus hauts scores indiquent une flexibilité psychologique plus importante face au travail. Dans notre étude, la WAAQ a été administrée sur la plateforme Limesurvey de l'IRPPSY, et a été utilisée pour évaluer la relation entre la flexibilité psychologique et les résultats psychothérapeutiques, pour étudier comment la relation des psychothérapeutes avec leur propres émotions pouvait affecter leurs résultats. Par ailleurs, dans son travail doctoral d'étude des « super-thérapeutes », notre collègue Daryl Chow (2014) avait étudié le lien entre le concept de « Disposition fixée de l'esprit » ou de « Disposition de l'esprit/mentalité de croissance » (Fixed Mindset vs Growth Mindset) issu des travaux de Carol Dweck, sans pouvoir mettre en évidence un lien entre ce construit et les résultats psychothérapeutiques. Nous faisons l'hypothèse que la capacité de flexibilité psychologique, définie comme l'habileté à engager des actions même en présence de pensées et de ressentis

difficiles, en direction d'objectifs valorisés, puisse être liée, ici significativement, à la capacité des thérapeutes à être plus efficaces ; en relation avec leurs capacités à détecter les émotions. En effet, les thérapeutes les plus à même de détecter des émotions négatives sont plus enclins à ressentir ces émotions, et leur capacité à agir, même en présence de contenu ressenti négatifs, pourrait en ce sens jouer un rôle sur les résultats attendus au traitement.

1.6.3.5. Le questionnaire Mini Profile of Nonverbal Sensitivity

Ce questionnaire (MiniPONS ; Bänziger, Scherer, Hall, Rosenthal, 2011), est une échelle auto-administrée, composée de 64 questions affichées à la suite du visionnage de 64 extraits vidéo d'une jeune femme illustrant différentes situations. A chaque présentation vidéo, il est proposé de choisir entre deux options de réponses décrivant la situation avec une seule réponse juste. Ce test mesure les capacités des sujets à reconnaître les signaux communiquant les ressentis, les attitudes et les intentions, à partir des expressions non-verbales véhiculées au travers de différents canaux de communication (dans la voix, dans les gestes, et dans la posture). Ce test comprend 32 vidéoclips qui sont non-sonorisés (16 vidéoclips où le visage est filmé et 16 vidéoclips où le corps est filmé) ; 16 vidéoclips qui contiennent uniquement des enregistrements audio, avec soit un contenu verbal qui est filtré - content filtered ou cf- soit un contenu verbal qui a été découpé et réorganisé de manière randomisée temporellement pour éliminer le contenu signifiant verbal- randomized spliced ou rs ; et 16 vidéoclips où le visage est affiché avec le son (Visage&Son). Ce test mesure donc la facette perceptive de la sensibilité interpersonnelle (Hall & Bernieri, 2001). L'administration de ces vidéoclips correspond en outre à quatre grandes catégories de situations ou des signaux non verbaux sont véhiculés (Bänziger et al, 2011) : Positive-Submissive « Positif-Soumis » (PS), les auteurs citent par exemple la situation de démonstration de la gratitude. Positive- Dismissive- pour « Positif dominant » (PD) par exemple, exprimer de l'affection maternelle ; Negative-Submissive ou « Négatif-Soumis » (NS) par exemple, demander pardon ; et Negative Dominant pour « Négatif-dominant » (ND), par exemple, la situation de menacer quelqu'un.

Le score total à la MiniPons est obtenu par l'addition des scores aux différents quadrants, il est compris entre 0 et 64 ; de meilleurs scores indiquent de meilleures capacités perceptives dans la sensibilité interpersonnelle. Nous avons utilisé la Mini-Pons dans cette étude pour évaluer les relations entre, la sensibilité non verbale des thérapeutes dans sa globalité et par canaux, et les résultats obtenus à la psychothérapie.

1.6.3.6. Le questionnaire Geneva Emotion Recognition Test Short Version

La GERT-S (Schlegel, Grandjean, & Scherer, 2014 ; Schlegel & Scherer, 2016) est une échelle auto-administrée, composée de 42 items, qui constitue un test de performance destiné à évaluer la capacité à reconnaître les émotions. 14 émotions sont « jouées » par des acteurs, et les émotions sont reproduites avec différents canaux (le visage, la voix et le corps). Ainsi, la présentation des émotions est plus écologique car elle est multimodale et dynamique. Le score total à la GERT-S est obtenu en additionnant les scores aux différents items codant pour chaque émotion (avec un score compris entre 0 et 3 pour chaque émotion) pour un score total compris entre 0 et 42 ; avec des plus hauts scores qui sont l'expression de meilleures capacités à reconnaître les émotions. Dans cette étude, nous avons utilisé la GERT-S pour évaluer la capacité des thérapeutes à reconnaître les macro-émotions en relation avec la taille de l'effet des patients qu'ils ont traités.

1.6.3.7. La tâche de reconnaissance des Micro-expressions faciales émotionnelles (MIE-R)

Il s'agit d'une tâche que nous avons construite spécifiquement pour cette recherche en sélectionnant des courtes vidéos de micro-expressions fiables, capturées sur le vif, et obtenues dans des conditions de laboratoire et qui ont été identifiées comme valides en utilisant le système FACS (Ekman, Friesen & Hager 2002). Cette base de micro-expressions « réelles » obtenue à partir des données de 32 participants, a été créée par Davison et al. (2016) à l'aide d'une tâche de provocation (de délivrance de stimuli à contenus émotionnels) et donc destinés

à susciter des émotions spécifiques tout en demandant au sujet de tenter de ne pas laisser transparaître d'émotion sur son visage, afin de faciliter l'apparition de micro-expressions. Les données enregistrées ont été ensuite classifiées par des codeurs FACS certifiés. Nous avons obtenu l'autorisation du docteur Davison d'utiliser certains clips pour créer notre tâche, et de les présenter à l'aide d'un questionnaire accessible sur internet, à partir d'une plateforme Limesurvey opérant à partir de notre site internet ; et validé avec lui que notre solution technique avec un accès sécurisé par mot de passe, était adéquate pour protéger ce matériel pour les participants de son étude, et pour le contenu en lui-même. Dans le consentement éclairé des thérapeutes il était explicité que le matériel compris dans le questionnaire était soumis à copyright et ne pouvait pas faire l'objet d'un quelconque enregistrement.

Nous nous sommes initialement intéressés aux travaux du Dr Paul Ekman et à la classification FACS dans le courant des années 2000, plus précisément entre 2005 et 2010, et nous avons été initialement formés à la détection des micro-expressions (FACE Training, février 2011). Nous avons utilisé la classification issue du manuel Facial Action Coding System depuis pour nous auto-former à la détection des Unités d'Action (AU) qui permettent de coder les expressions faciales, et qui permettent également de définir des combinaisons d'AU révélant la présence des sept émotions primaires. Il était ainsi initialement convenu d'utiliser le test de reconnaissance des micro-expressions administrable en ligne développé par Paul Ekman (METT), dans le cadre d'une collaboration, de le traduire en français, et de tester si les différences de capacités dans la reconnaissance des micro-expressions pouvaient être liées à de meilleurs résultats en psychothérapie en lien avec la pratique FIT.

Faisant face à l'impossibilité effective d'utiliser ce test, nous avons dû trouver une autre solution technique et avons effectué une nouvelle revue de la littérature sur l'évaluation des habiletés non-verbales. Nous avons dès lors voulu construire une tâche de performance qui puisse dépasser les limitations des approches disponibles d'évaluation de la capacité à décoder les micro-expressions faciales émotionnelles.

En effet, les tests disponibles (notamment ceux présentés par le Dr Ekman) sont considérés comme des standards dans l'évaluation des capacités à détecter les micro-expressions, et les formations ou l'entraînement à ces habiletés reposent en partie sur ces tests pour évaluer un niveau de base de performance ou pour suivre la progression (METT training). Pourtant, ces tests utilisent des acteurs qui reproduisent des Unités d'Action d'expressions émotionnelles à la demande de l'auteur du test. Avec, logiquement, un effet d'augmentation de la consistance interne pour les différentes catégories émotionnelles, puisque pour chaque émotion, un nombre égal de présentation avec des sujets différents (Homme, Femme, groupes démographiques différents) est présenté (par exemple quatre questions/affichages par libellé d'émotion), avec les mêmes Unités d'Action qui sont reproduites sur commande. Pourtant, dans la classification FACS, il est mis en avant que l'expression des émotions puisse nécessiter l'activation de certains muscles faciaux dits fiables, mais aussi qu'il puisse exister des variantes dans l'expression d'une même émotion, avec certaines Unités d'Actions présentes de manière facultative. Utiliser la même combinaison d'Unité d'Action produite sur commande pourrait donc limiter l'étude des variantes émotionnelles et de leur expression naturelle et serait donc moins écologique, bien que cette approche puisse permettre de créer une « vraie » échelle de mesure à partir de la tâche de performance (il est notamment nécessaire que les items d'une catégorie mesurent peu ou prou le construit d'une manière similaire, c'est-à-dire théoriquement qu'ils soient *tau* équivalents, pour autoriser le calcul d'un coefficient de consistance interne).

Des interrogations sur les limitations méthodologiques dans l'étude des micro-expressions, sont notamment retrouvées dans l'article de Li et al. (2015) sur la création d'un système automatisé de reconnaissance des micro-expressions : certains systèmes informatiques pourraient notamment obtenir de très hauts scores de reconnaissance des micro-expressions avec la base METT parce que les extraits de micro-expressions contenus dans la base METT sont « entièrement synthétiques » ; puisqu'ils sont constitués d'une image d'expression faciale émotionnelle, insérée dans une séquence d'images neutres, rendant la détection

significativement plus aisée. En condition naturelle l'apparition des micro-expressions n'étant pas aussi tranchée et le contexte étant plus dynamique.

En effet, afin de réduire les effets sur la reconnaissance des émotions dus au visionnage de la dernière image véhiculant une émotion, et qui tranche nettement par rapport à la première image dans les clips vidéos, les auteurs de ces tests utilisent classiquement une approche en « sandwich » de la présentation des expressions qui est à même de favoriser une meilleure reconnaissance : c'est-à-dire que l'expression faciale émotionnelle, qui n'est pas présentée d'ailleurs directement au format vidéo, ce qui représente une transformation supplémentaire du signal, est présentée comme une sorte de GIF animé en étant précédée et suivie de l'expression faciale neutre du sujet dont l'expression émotionnelle a été initialement enregistrée. Ainsi, il n'y a plus de persistance possible dans l'affichage de la dernière image de l'expression émotionnelle, qui, à la manière d'une tâche de mémorisation et de rappel de mots où la liste n'est pas continue mais finie, pourrait perturber la reconnaissance (puisque dans le cas d'un rappel de mots le dernier mot de la liste est accessible en mémoire à court terme) ; ici, la dernière image affichée étant neutre, le stimulus est moins saillant. Mais, nous pouvons observer que, bien sûr, en conditions naturelles, les expressions faciales émotionnelles n'existent pas et ne sont pas affichées sous forme de « GIF » et que le test des micro-expressions du METT, pour caricaturer l'expression de Binet sur l'évaluation du quotient intellectuel, ne mesure pas forcément l'habileté d'un sujet à détecter une micro-expression telle qu'elle apparaît naturellement dans une interaction, mais mesure la capacité à reconnaître différents GIF reproduisant une variante prototypique d'une micro-expression posée par des acteurs entraînés.

Afin de dépasser ces limitations, nous avons recherché les bases disponibles et utilisables pour des recherches qui pourraient contenir de « vraies » micro-expressions. C'est-à-dire des enregistrements où ces expressions ne seraient pas posées, mais seraient présentes spontanément (ce ne sont pas des acteurs qui les produisent à la demande), et où les différentes variantes des émotions seraient classifiées en fonction des unités d'actions émotionnelles par

des experts dans la codification FACS, tout en contenant les 7 émotions primaires. Nous n'avons identifié à ce jour qu'une base qui satisfasse à tous ces critères, la base créée par Davison et al. (2016).

Construction de la tâche de performance.

Les extraits contenant des micro-expression ayant été découpés et fournis au format « image » par l'équipe de chercheur, nous avons reconstruit les extraits vidéos en respectant les conditions initiales d'enregistrement, pour maintenir la durée d'affichage de chaque extrait vidéo et sa qualité, et validé ces options avec le Dr Davison. Ainsi, nous avons utilisé le logiciel Matlab (R2016b) pour effectuer un rendu vidéo avec les paramètres suivants (*qualité de la vidéo = 75 %*, *nombre d'images par secondes = 200*).

Nous avons donc commencé par sélectionner dans la base, les clips vidéo qui étaient classifiés comme appartenant aux sept micro-expressions cibles, et nous avons vérifié que les durées de présentation de ces expressions étaient bien inférieures au seuil classifiant pour une micro-expression, qui n'est toutefois pas l'objet d'un fort consensus (avec un seuil équivalent à une durée $< \frac{1}{2}$ s), nous avons vérifié que les premières images ne contenaient pas de micro-expressions parasites auprès de l'auteur, puis sélectionnés pour chaque émotion cible quatre clips vidéos représentatifs, en essayant d'assurer une progressivité de la difficulté de reconnaissance et une représentativité égale entre les sexes (ce résultat ne put pas être atteint puisque nous ne disposions pas d'un nombre équivalent de vidéoclips représentant chacune des émotions pour chaque sexe). Une fois que chaque émotion était associée à quatre clips distincts, nous avons envoyé les clips sur le serveur sécurisé Limesurvey de notre site internet expérimental.

Nous avons paramétré l'affichage des vidéos produit par le logiciel dans le navigateur internet en utilisant un code javascript pour présenter un masque (un symbole carré avec une bordure) pendant le chargement de la vidéo sur internet. En effet, lors de notre phase de pré-test de la tâche impliquant 3 sujets avec des ordinateurs différents et des navigateurs différents,

nous avons constaté que suivant la connexion internet, les vidéos pouvaient se figer ou s'afficher plus ou moins rapidement une fois qu'elles étaient chargées par le navigateur, ce qui était évidemment non acceptable face à la nécessaire standardisation de la tâche. De même la vidéo pouvait pendant le chargement être affichée comme une image, avec la première image de la séquence vidéo apparaissant figée, et présentant une expression faciale neutre ou différente de la micro-expression à détecter, un peu à la manière de la présentation du test METT, ce qui favorisait grandement la détection (« spotting ») de la micro-expression. Aussi, nous avons utilisé et modifié un code javascript fourni et utilisé par le Professeur Marc Covents dans l'administration de l'échelle GERT-S, pour forcer le chargement complet de la vidéo avant son affichage par le navigateur. Ainsi chaque participant peut regarder la vidéo dans les mêmes conditions. Pour ce faire, nous avons forcé le navigateur à afficher un symbole carré, pendant le chargement de la page, qui agit à la manière d'un masque utilisé dans les expériences de psychologie cognitive.

Cette solution est bien entendue imparfaite, puisque selon la vitesse de connexion, certains participants, auront un affichage de ce symbole-masque un peu plus long, mais elle permet de coller au plus près des conditions de laboratoire, que permettraient l'utilisation de la présentation vidéo sur site avec un poste de travail dédié utilisant un logiciel de présentation avec un horodatage (comme le logiciel E-Prime). Il est à noter qu'il n'est pas possible dans ces conditions de s'assurer de la taille de l'écran qui varie d'un utilisateur à l'autre. Par contre, nous avons forcé l'utilisation d'un ordinateur et non pas d'une tablette pour standardiser au mieux la passation, et l'utilisation du navigateur Firefox (fondation Mozilla), qui reproduisait le plus fidèlement notre tâche. Nous avons également produit une brochure expliquant notre protocole aux thérapeutes et une procédure à suivre en cas de difficultés.

Bien que notre pré-test et notre crash test de la plateforme rendent essentiellement des informations de nature qualitatives, et que notre premier test d'affichage des micro-expressions ait été quelque peu biaisé du fait de la présentation initiale d'images figées (« effet poster »),

nous avons pu observer que la performance au test était croissante entre les trois sujets (sujet non sensibilisé aux micro-expressions, sujet sensibilisé et non entraîné, sujet sensibilisé et entraîné).

1.6.3.8. Variables contextuelles et de contrôle collectées au niveau du patient.

Nous avons obtenu à partir de l'extraction de la base de données des thérapeutes et quand ces données étaient disponibles, les variables âge, sexe, et le nombre de séance de thérapie, la ou les mesures de résultats à la première séance sur les échelles PHQ-9, GAD7 ou CORE-10 (entrée dans le traitement (T1), la ou les mesures de résultats à la dernière séance enregistrée avec ces échelles (T2). Les données de diagnostic des patients n'étaient pas disponibles.

1.6.3.9. Variables contextuelles collectées au niveau des thérapeutes

Nous avons enregistré dans notre enquête les variables âge, sexe, profession, niveau de diplôme, orientation thérapeutique, expérience clinique en années, nombre d'heures de supervision suivie sur les trois derniers mois, et nombre de jours de formation sur l'année précédente.

Nous avons interrogé les thérapeutes sur la quantité de temps dévolue par chacun à la pratique délibérée sur le dernier mois, pour éviter un biais de rétrospection, plutôt que de les interroger sur leur quantité de pratique délibérée réalisée sur l'année précédente ; nous avons également demandé aux thérapeutes - du fait de notre design rétrospectif - et à des fins de contrôle si cet agenda d'entraînement était un agenda représentatif de leur entraînement habituel à la pratique délibérée. La dimension de pratique délibérée interrogée renvoyait à la question spécifique formulée par notre collègue Daryl Chow (Chow et al., 2015) dans son travail princeps sur l'impact de la pratique délibérée sur l'efficacité des psychothérapeutes : « Combien d'heures avez-vous passé seul, fortement engagé dans des activités reliées à l'amélioration des vos capacités de thérapeute ». Nous avons enfin contrôlé si les thérapeutes

avaient suivi des formations à la reconnaissance des émotions sur les trois années précédentes et en particulier des formations à la reconnaissance des micro-expressions. En effet, du fait de notre design rétrospectif, il était également nécessaire de s'assurer qu'aucun entraînement formel n'avait pu impacter la performance des thérapeutes dans la reconnaissance des signaux non-verbaux qui est, en dehors de ces conditions d'entraînement, réputée suffisamment stable.

1.6.3.10. Analyses statistiques

Les traitements statistiques ont été réalisés avec le logiciel SPSS (V25) et le logiciel SAS (9.04) pour l'analyse Multiniveau. Il s'agit d'une méthodologie qui permet de traiter des données pour lesquelles, le contexte exerce une importance significative. Dans les études cliniques des effets dus aux psychothérapies, il existe souvent une relation de dépendance entre les données recueillies et les contextes de soins. Les patients de l'étude de Anderson et al. (2009) par exemple, ont été traités par vingt-cinq thérapeutes, et à chaque thérapeute était associé un groupe de patient, créant une relation de dépendance entre le thérapeute et les patients (ses données). Intuitivement, les patients qui sont traités par un même thérapeute sont censés partager plus de caractéristiques communes et donc à être plus semblables à leurs homologues traités par le même thérapeute (en termes de réception du traitement et donc de résultats), qu'à ceux traités par un des vingt-quatre autres thérapeutes de l'étude. Ces patients appartiennent donc à un contexte de soins, ou à un niveau supérieur d'abstraction, à un groupe qui est le groupe défini par leur thérapeute particulier. Il est ainsi possible de distinguer un premier niveau inférieur dans les données, le niveau des données du patient, où il est possible de procéder à une étude des différences interindividuelles. Et comme ces patients étaient traités par un des vingt-cinq thérapeutes respectifs, leurs données individuelles de ce premier niveau sont donc « nichées » (de l'anglais « nested ») et reliées au contexte de soins de leur thérapeute, qui représente un deuxième niveau. La relation de dépendance hiérarchique des données récoltées au niveau des patients (niveau 1) en relation à leur thérapeute à un niveau supérieur (niveau 2) implique une plus grande prudence dans l'analyse de la différence des données

obtenues par les patients au niveau 1, comme par exemple leur réussite au traitement ou le changement thérapeutique constaté, puisque les données de chaque patient traité par le même thérapeute ne sont pas de ce fait totalement indépendantes.

L'analyse multiniveau répond à ces enjeux en permettant de partitionner la variance des résultats entre les données qui sont récoltées au niveau des patients et donc attribuables aux patients et les données attribuables aux contextes de soins auxquels appartiennent ces patients (les différents thérapeutes). De manière simplifiée, s'il existe plus de variance des résultats entre les données de chaque patient qu'entre les données de chaque groupe de patients défini par le même thérapeute, le contexte de soins (thérapeute) semble logiquement moins bien expliquer la différence des résultats que les différences dues aux patients considérés séparément ; et le niveau thérapeute n'influence vraisemblablement pas les données : ces données pourraient être analysées classiquement comme si les patients étaient indépendants, c'est-à-dire au niveau des données des patients eux-mêmes sans tenir compte de l'influence des variables liées au thérapeute. Dans le cas contraire, le contexte « thérapeute » explique une partie significative des différences de résultats observées au traitement et il s'agira de traiter les données en analysant les variables au niveau des thérapeutes, ce que permet de faire l'analyse multiniveau.

1.6.4. Résultats

Afin d'éprouver nos hypothèses, nous avons réalisé différents traitements statistiques.

Pour mémoire, notre hypothèse H2 postulait que la capacité à reconnaître des signaux non verbaux par le biais de différents canaux (corps, visage, voix) pouvait être associée aux résultats de la psychothérapie. Notre hypothèse H3 postulait que la capacité à reconnaître les micro expressions était significativement plus corrélée avec les résultats de la psychothérapie que la capacité à reconnaître les macro expressions. Notre hypothèse H4 postulait que la capacité à reconnaître les micro expressions négatives étaient plus spécifiquement associée aux résultats du traitement.

Enfin, notre hypothèse H5, postulait qu'en condition de feedback instrumental, la capacité à décoder les signaux non verbaux était associée à de meilleurs résultats amenant un premier argument à l'hypothèse supplétive du feedback instrumental.

Nous avons réalisé une analyse de corrélation entre les scores obtenus par les thérapeutes aux tâches de reconnaissance des signaux non verbaux dans la voix, le visage et le corps (en macro et en micro) et entre les résultats obtenus au traitement pour tester ces différentes hypothèses. En effet, trouver une corrélation significative entre les résultats au traitement et les capacités des thérapeutes allait dans le sens de l'hypothèse H1, H2 et H5.

Trouver une taille de l'effet plus importante entre la reconnaissance des micro-expressions et les résultats (en comparaison des macro expressions) apportait des éléments de réponse pour l'hypothèse H3.

Nous avons donc analysé ces résultats globaux pour répondre à ces hypothèses particulières.

Nous avons d'abord à des fins exploratoires, et bien que cette analyse soit moins fiable, étudié ces corrélations au niveau des scores des patients (où les mesures des capacités des thérapeutes sont répétées puisque les patients « partagent » pour certains le même thérapeute), puis, au niveau des scores de l'échantillon de thérapeutes, à l'aide de la modélisation multiniveau. Enfin, nous avons corrigé ces analyses en tenant compte des seuils de sévérité, c'est à dire de la détresse psychologique, présentés par les patients traités dans cet échantillon.

1.6.4.1. Statistiques descriptives concernant les thérapeutes

Les 11 thérapeutes de cette étude (7 femmes, 4 hommes) étaient âgés de 42.18 ans en moyenne ($ET = 8.05$). 5 praticiens se sont déclarés comme étant des psychothérapeutes, un praticien exerçait le counseling, et 5 autres praticiens appartenaient à la catégorie « autres » (n'étant ni définis comme psychiatres, psychologues, exerçant le counseling ou comme travailleurs sociaux). A ce sujet, il est nécessaire de rappeler qu'il n'a pas été retrouvé dans la

littérature de différence d'efficacité des psychothérapies entre les différents corps de métier exerçant la relation d'aide, résultats qui a été répliqués et retrouvé dans l'étude contrôlée randomisée d'Anderson que nous avons citée dans notre introduction.

En termes de niveau d'éducation, quatre des thérapeutes ont déclaré se situer à un niveau d'éducation correspondant à un diplôme (« Graduate »), 3 ont déclaré se situer au-dessus de ce niveau (« post graduate ») et 4 ont déclaré qu'ils n'avaient pas de diplôme spécifique (« degree »).

En termes d'orientation de la pratique, 5 thérapeutes se sont déclarés éclectiques-intégratifs (pratiquant une combinaison de thérapies suivantes : thérapie psychodynamique, TCC, humaniste/centré sur le client, thérapie familiale et systémique, thérapie stratégique / modèle de Palo-alto, thérapie Eriksonienne/hypnose, thérapie existentielle), 2 ont déclaré pratiquer la thérapie Ericksonienne et l'hypnose, et 4 se sont situés dans la catégorie autre (une autre forme de thérapie que la thérapie psychodynamique, que la TCC ou la thérapie humaniste/centrée sur le client, thérapie familiale et systémique, thérapie stratégique / modèle de Palo-alto, thérapie eriksonienne/hypnose, ou la thérapie existentielle).

L'expérience moyenne des thérapeutes était de 9.73 ans ($ET = 6.07$). Le nombre d'heures de supervision reçu sur les 3 derniers mois était compris entre 0 et 6 heures ($M = 2.55$; $ET = 1.57$). Les thérapeutes avaient reçu en moyenne 2.64 jours de formation sur l'année précédente ($ET = 2.46$). Les scores des thérapeutes quant à la flexibilité psychologique au travail, étaient de 25 à 48 avec une moyenne de 35.73 ($ET = 6.36$).

Les thérapeutes ont été interrogés également sur leur « pratique délibérée » au cours du mois précédent : ils avaient dévolu un temps moyen de 13.96 heures à cette pratique avec une grande disparité dans leurs réponses ($ET = 11.46$ h, *minimum* = 3, *maximum* = 30). Six thérapeutes ont répondu que cette quantité de pratique délibérée était typique de leur agenda et cinq ont affirmé le contraire.

Aucun des thérapeutes n'avait suivi de formation à la reconnaissance des émotions en général, et à la reconnaissance des micro-expressions en particulier, sur les trois dernières années ce qui laissait à supposer que leurs habiletés à détecter les émotions ou les signaux non-verbaux était suffisamment stable dans le temps.

Ces onze thérapeutes possédaient un registre de consultations pour lesquelles les scores à l'entrée dans les soins et à la dernière séance étaient disponibles. Dans ces données chaque thérapeute avait suivi entre 11 et 91 patients.

Ces variables, à l'exception de la quantité de pratique délibérée, et du score de flexibilité psychologique, ne seront pas utilisées dans nos analyses statistiques et sont reproduites ici pour illustrer le contexte de notre étude. En effet, les variables orientation du traitement, expérience des cliniciens, et filière professionnelle, ne sont pas identifiés dans la littérature comme des prédicteurs associés aux résultats des traitements de manière consistante (Bertolino et al., 2012).

1.6.4.2. Statistiques descriptives concernant les patients.

Le jeu de données initial mis à notre disposition contenait les données de 534 patients pour les 11 thérapeutes qui ont pris part à notre protocole à T1. Nous n'avons pu conserver que les enregistrements permettant de calculer une taille de l'effet et donc les enregistrements où étaient présents les scores aux questionnaires à la fois à la séance initiale et à la séance finale (dernière séance disponible). Après vérification et élimination des données manquantes, le jeu final de données fut constitué de 386 patients (229 femmes, 144 hommes et 13 patients de sexe indéterminé). L'information sur l'âge était disponible pour 354 patients et manquante dans 32 cas. L'âge moyen des patients était de 40.79 ans ($ET = 11.83$; avec des patients âgés de 16 à 79 ans). Les 386 patients de cette étude ont reçu un nombre minimal de 2 séances de psychothérapie et un nombre maximal de 31 séances ($M = 4.60$; $ET = 2.82$). L'arrêt du traitement dans cet échantillon était attendu dans 66.6 % des cas et correspondrait donc à un abandon dans 33.4 % des cas.

1.6.4.3. Données de résultats mesurées pendant les soins

215 patients ont rempli le questionnaires CORE-10 à T1 et T2 (Tableau 8), et 171 patients les questionnaires PHQ-9 et GAD-7 sur ces deux occasions de mesure (Tableaux 9 et 10).

Tableau 8. Scores bruts obtenus à l'échelle CORE-10 à la première séance (T1) et à la dernière séance de thérapie disponible (T2 ; $n = 215$)

	T1	T2
Moyenne	18.44	9.86
Médiane	19.00	8.00
Écart type	7.61	7.38
Score minimum	0	0
Score maximum	36	34

Tableau 9. Scores bruts obtenus à l'échelle GAD-7 à la première séance (T1) et à la dernière séance de thérapie disponible (T2 ; $n = 171$)

	T1	T2
Moyenne	9.69	5.07
Médiane	9	4.00
Écart type	5.65	4.24
Score minimum	0	0
Score maximum	21	21

Tableau 10. Scores obtenus à l'échelle PHQ-9 à la première séance (T1) et à la dernière séance de thérapie disponible (T2 ; $n = 171$)

	T1	T2
Moyenne	9.10	5.11
Médiane	8.00	4.00
Écart type	5.72	4.65
Score minimum	0	0
Score maximum	26	26

1.6.4.4. Données de résultats relatives aux habiletés des thérapeutes

Capacité à détecter les macro-manifestations d'expressions émotionnelles

Tableau 11. Scores bruts globaux obtenus à la reconnaissance des macro-émotions (GERT-S, $n = 11$).

	Total	Total Émotions Négatives	Total Émotions Positives
Moyenne	28.91	13.27	14.09
Écart type	4.81	3.23	2.47
Minimum	23	8	10
Maximum	38	18	18

Les thérapeutes ont sensiblement identifié les émotions négatives et positives avec la même efficacité (Tableau 11). Plus spécifiquement, ils ont plus facilement discriminé les émotions à connotations positives que sont le soulagement, l'intérêt, le plaisir, et la fierté et certaines émotions négatives (la colère, l'irritation, le désespoir, et le dégoût) que l'anxiété, la

joie, la surprise. La tristesse et la surprise apparaissent comme les émotions les plus difficiles à discriminer, pour cet échantillon de thérapeutes (Tableaux 12 a et b).

Tableaux 12. a) et 12. b) des scores bruts obtenus à la reconnaissance des macro-émotions, par émotion (GERT-S, $n = 11$).

	Colère	Fierté	Joie	Amusement	Plaisir	Soulagement	Intérêt
Moyenne	2.09	2.36	1.82	1.91	2.55	2.82	2.64
Écart type	.83	.67	.98	.94	.69	.41	.51
Minimum	1	1	0	0	1	2	2
Maximum	3	3	3	3	3	3	3

Tableau 12 a)

	Surprise	Anxiété	Peur	Désespoir	Tristesse	Dégoût	Irritation
Moyenne	1.55	1.91	2	2.09	1.09	2	2.09
Écart type	.93	.94	.89	.83	1.04	.63	.70
Minimum	0	0	1	1	0	1	1
Maximum	3	3	3	3	3	3	3

Tableau 12 b)

Capacité à détecter les signaux non verbaux (MINI-PONS)

Comme nous avons rencontré des difficultés techniques lors de la passation du questionnaire Mini-Pons, un bug informatique ayant entraîné la duplication de certaines questions et les informations ne pouvant pas être récupérées du côté de l'université administrant le questionnaire en Suisse, nous avons directement calculé pour chaque modalité un score moyen obtenu par le score total de la modalité, divisé par le nombre d'items réellement cotés.

Le score moyen est donc compris entre 0 et 1. Le tableau 13 reprend la moyenne de ces résultats pour l'ensemble des 11 thérapeutes.

Tableau 13. Scores moyens obtenus à la reconnaissance des signaux non verbaux par cadrants et modalités (MINI-PONS, $n = 11$).

	Cadrant PS	Cadrant NS	Cadrant PD	Cadrant ND	Score Total
Moyenne	.76	.77	.79	.82	.79
Écart-Type	.10	.09	.10	.09	.06
Minimum	.63	.67	.63	.69	.69
Maximum	.94	.94	.88	1	.88

Notes :PS = Positif Soumis, NS = Négatif Soumis, PD = Positif Dominant, ND = Négatif Dominant

Le score total moyen à la MINI-PONS est assez élevé sur notre échantillon de thérapeutes (79 % de bonnes réponses) avec plus de facilité dans la reconnaissance des quadrants Positive Dismisive et Negative dismissive.

Nous pouvons observer que sur notre échantillon de onze thérapeutes, les signaux véhiculés par la combinaison visage et voix, dont le contenu a été filtré, ont été plus facilement reconnus que les signaux véhiculés par le visage et la voix, dont la séquence a été découpée et réarrangée aléatoirement. Les signaux issus de la voix dont le contenu a été filtré et ceux issus du canal « corps » seul, ont été les plus difficiles à identifier (tableaux 14 a et b).

Tableau 14. a) 14. b) Scores moyens obtenus à la reconnaissance des signaux non verbaux par modalités (MINI-PONS, $n = 11$).

	Visage	Corps	Voix CF	Voix RS	Visage&Voix CF	Visage et VoixRS
Moyenne	.80	.71	.75	.80	.92	.81
Écart type	.11	.09	.10	.13	.06	.14
Minimum	.69	.56	.63	.63	.88	.63
Maximum	1.00	.88	.88	1	1	1

Tableau 14. a)

	Pas de son (Visage et corps)	Voix seule	Voix & Visage
Moyenne	.75	.77	.87
Écart type	.07	.08	.07
Minimum	.66	.69	.75
Maximum	.84	.94	.94

Tableau 14. b)

Capacité à détecter les micro-manifestations d'expressions faciales émotionnelles

(tâche MIE-R)

Tableau 15. Scores bruts obtenus à la reconnaissance des micro expressions (MIE-R, $n = 11$)

	Total	TN	TP (Joie)	Su	Co	Pe	Tr	Dé	Mé
Moyenne	5.73	2.82	1.82	1.09	.45	.36	.73	.64	.64
Écart-type	1.74	1.33	.41	1.22	1.04	.67	.47	.67	.81
Minimum	3	1	1	0	0	0	0	0	0
Maximum	8	5	2	4	3	2	1	2	2

Note. TN = Score brut total de détection des micro-émotions négatives, TP = Score brut total de détection des micro-émotions positives. Su = surprise, Co = colère, Pe = peur, Tr = tristesse, Dé = dégoût, Mé = mépris

Nous pouvons observer que le score total moyen de reconnaissance des micro-expressions est bas (tableau 15 ; 5.73 sur 28) avec un intervalle des scores réduit, qui signale un effet plancher pour notre échantillon testé. Cette tâche paraît donc extrêmement difficile. Nous n'avons pas pu contrôler les conditions de passation (tranquillité, concentration) qui auraient pu en outre affecter les résultats de nos sujets. Ce résultat est toutefois théoriquement attendu puisque sans entraînement formel, peu d'individus sont capables de reconnaître les micro-expressions, notamment dans la détection du mensonge (Ekman, O'Sullivan & Frank, 1999). Ceci entraîne que la distribution des scores sur la population ne puisse pas suivre une distribution normale, avec une mince portion de la population qui pourrait obtenir des scores élevés, et la majorité des individus qui obtiendraient des scores très bas.

Les micro-émotions négatives ont été moins facilement discriminées sur notre échantillon (moyenne de la catégorie = .47) que la seule émotion positive (joie ; $M = 1.82$). Par ordre décroissant de meilleure reconnaissance, la micro-expression de surprise est la micro-expression la mieux reconnue, avec un de nos thérapeutes qui a reconnu tous les items ($M = 1.09$), puis suivent la tristesse ($M = .73$), le dégoût et le mépris ($M = .64$), puis enfin la colère ($M = .45$) et la peur ($M = .36$) qui sont les micro-émotions les moins facilement distinguées ici.

1.6.4.5. Analyse statistique des relations entre les capacités des thérapeutes à reconnaître les signaux non verbaux et les performances de thérapeutes.

Pour pouvoir analyser nos données nous avons utilisé un script produit par Jebb Brown et Tak Minami du Center for Clinical Informatics (CCI- états Unis). En effet, notre but était de pouvoir analyser des relations entre nos variables d'intérêt (les différentes capacités à reconnaître les signaux non verbaux) et les résultats obtenus par les thérapeutes avec leurs patients. Ces résultats ayant été mesurés avec trois questionnaires différents, le questionnaire CORE-10, le questionnaire PHQ-9, et le questionnaire GAD-7 ; il était possible d'analyser les résultats séparément mais nous aurions perdu de la puissance statistique, c'est-à-dire de la capacité à détecter un effet, s'il existe, si nous ne pouvions pas tester les relations entre nos

variables d'intérêt sur notre population considérée dans son ensemble. Et pour pouvoir effectuer ces traitements, il était nécessaire de transformer nos scores de résultats en une métrique commune, une taille de l'effet, pour pouvoir utiliser un seul indice de résultat par sujet de l'étude. Ceci a été rendu possible par le fait que ces différents questionnaires, même s'ils mesurent des variables différentes, mesurent en fait une même métrique commune sous-jacente de détresse psychologique. L'équipe du Dr Jebb Brown analyse ainsi ces questionnaires, dans son travail auprès des cliniciens et des institutions en agrégeant les scores obtenus au PHQ-9 et au GAD-7.

Obtention d'une métrique commune de résultats :

Calcul des scores moyens

Nous avons utilisé le logiciel SAS pour tout d'abord transformer les scores bruts obtenus aux différents questionnaires, en calculant un score moyen (en divisant le score brut par le nombre d'items administré à chaque échelle). Pour l'échelle PHQ-9 et pour la GAD-7, nous avons ensuite obtenu un score combiné PHQGAD en calculant une moyenne pondérée par le nombre d'item à chaque échelle.

Calcul des scores standardisés

Nous avons ensuite calculé pour chaque questionnaire, les scores moyens et l'écart-type obtenus à l'entrée dans le traitement (T1). Puis transformé les scores moyens obtenus pour chaque questionnaire, à T1 et T2 en scores standardisés Z , en utilisant ces données de moyenne et d'écart-type obtenues en T1. En effet, nous avons utilisé les scores de référence (M et ET) obtenus en T1 pour calculer le score Z en T2, puisque nous souhaitons calculer un score de changement (taille de l'effet).

Calcul de la taille de l'effet standardisée (en unité de scores Z)

Nous avons ainsi pu calculer la taille de l'effet standardisée en effectuant une simple soustraction entre le score standardisé à l'occasion de mesurage numéro 1 (T1 ou entrée dans le traitement) et le score de la mesurage numéro 2 (T2 ou dernière séance disponible). Comme les trois questionnaires (CORE-10, PHQ-9, GAD-7) mesurent la détresse psychologique en sens direct, les scores attendus à T2 sont supposés inférieurs à ceux obtenus en T1, et la différence $Z1-Z2$ est la plus souvent attendue comme positive, chez les patients de cet échantillon.

Correction par les caractéristiques cliniques de l'échantillon

L'étape suivante consistait à réaliser une analyse de régression pour prédire les scores de changement standardisés¹⁴ au traitement par les scores des patients à l'entrée dans le traitement (T1). Cette opération nous a permis d'obtenir un score de changement standardisé corrigé par les résidus calculés lors de cette opération - Cette manipulation nous a en effet permis de corriger les tailles d'effet. Le score de changement « résidualisé » est en effet obtenu par la différence entre le score réellement observé à la fin du traitement chez le patient et le score de changement prédit par l'équation de régression, en tenant compte du score à l'entrée dans le traitement.

¹⁴ Des données complémentaires sur cette méthodologie peuvent être trouvées dans l'article suivant : Minami et al. (2012). Benchmarking therapists : furthering the benchmarking method in its application to clinical practice. *Quality and Quantity*, 46, 6, 1699-1708.

Calcul de l'effet du au thérapeute en utilisant la modélisation multiniveau

La modélisation multiniveau permet de calculer le pourcentage de variance des résultats (ici ajustés et standardisés) qui sont explicables par le (niveau) thérapeute. Ce pourcentage obtenu par le calcul du coefficient de corrélation intra-classe (ICC) est produit en calculant le paramètre estimé par le modèle d'analyse multiniveau pour le changement du au thérapeute, divisé par la somme des effets dus aux thérapeutes et à l'erreur estimée dans le modèle (paramètre / (paramètre +résidus)).

Ici nous considèrerons que les effets dus au thérapeute ne sont pas fixés car la population de thérapeute est seulement un échantillon issu de la population générale des thérapeutes, dont les effets peuvent varier, et ne peut dans ce cas la représenter parfaitement (random effects). Notre calcul a donné un ICC de .13, ce qui signifie que pratiquement 13% de la variance de la taille de l'effet due au traitement peut être attribuée aux différences entre les thérapeutes. Ce résultat diffère des résultats classiques obtenus dans les travaux de Wampold (Wampold & Imel, 2015) que nous avons déjà cités (aux alentours de 5 à 8%), mais il faut rappeler qu'ici notre échantillon de thérapeutes est très réduit.

1.6.4.6. Analyse exploratoire des corrélations entre les scores obtenus aux questionnaires de performance dans la reconnaissance des signaux non verbaux et les scores de changement dus au traitement.

Afin d'explorer les effets présents dans notre étude, nous allons commencer par analyser les matrices de corrélations entre les scores de résultats de nos thérapeutes aux tests de performance de reconnaissance des signaux non verbaux, et les scores de résultats obtenus auprès des patients qu'ils ont traités. Il est nécessaire d'interpréter ces résultats avec la plus grande prudence, puisque les différents patients ne sont pas indépendamment traités par des thérapeutes uniques, et les effets analysés peuvent donc être biaisés et augmentés du fait de cet artefact. Cette analyse est donc une première analyse « naïve » qui peut toutefois nous éclairer sur de possibles effets si leur magnitude est importante.

En analysant les relations entre les scores de changement des patients (Score de changement standardisés ou taille de l'effet, tableau 16) et les scores de reconnaissance des signaux non verbaux des thérapeutes qui leurs sont associés, nous observons pour l'échantillon total ($n = 386$) que la reconnaissance des macro-émotions dans leur ensemble, des macros-émotions positives, et plus spécifiquement des macro-émotions joie, plaisir, surprise, tristesse, et dégoût sont associées à des résultats positifs au traitement. Plus étonnamment, la capacité à détecter le soulagement, est reliée à de moins bons résultats.

En ce qui concerne la reconnaissance des micro-expressions, l'effet de la reconnaissance des micro-expressions négatives n'est pas significatif mais tendanciel (en sens positif), l'effet de la reconnaissance de la tristesse, de la peur et de la colère est significatif et positif. Étonnamment, l'effet de la reconnaissance de la surprise est négatif.

En termes de sensibilité non-verbale, la reconnaissance des informations présentes dans la voix avec un contenu filtré semble être reliée négativement vis-à-vis des résultats au traitement, celle de la reconnaissance des informations dans la voix au contenu RS est positive.

Tableau 16. Corrélations de Pearson (r) entre les scores de changement standardisés-ajustés obtenus auprès des patients et les scores de reconnaissance des signaux non verbaux obtenus par les thérapeutes. Échantillon Initial total ($n = 386$)

	Score de changement standardisé	
Reconnaissance Globale des macro-émotions	$r = .14$	$p = 0.01$
Reconnaissance des Macro émotions Positives	$r = 0.11$	$p = 0.03$
Reconnaissance Macro -Joie	$r = 0.18$	$p < 0.01$
Reconnaissance Macro-Plaisir	$r = 0.11$	$p = 0.04$
Reconnaissance Macro-Soulagement	$r = - 0.14$	$p < 0.01$
Reconnaissance Macro-surprise	$r = 0.11$	$p = 0.04$
Reconnaissance Macro-tristesse	$r = 0.15$	$p < 0.01$
Reconnaissance Macro-dégout	$r = 0.11$	$p = 0.04$
Reconnaissance Micro émotions Négatives	$r = 0.10$	$p = 0.06$
Reconnaissance Micro -surprise	$r = - 0.14$	$p < 0.01$
Reconnaissance Micro-tristesse	$r = 0.13$	$p < 0.01$
Reconnaissance Micro- peur	$r = 0.15$	$p < 0.01$
Reconnaissance Micro-dégout	$r = - 0.11$	$p = 0.03$
Reconnaissance Micro-colère	$r = 0.13$	$p < 0.01$
Reconnaissance Voix Contenu Filtré (CF)	$r = - 0.16$	$p < 0.01$
Reconnaissance Voix Contenu RS	$r = 0.18$	$p < 0.01$

Sous-analyse par questionnaire

Questionnaires CORE-10 combinés ($n = 215$)

Avec les scores obtenus par les patients au questionnaire CORE-10 (tableau 17), les relations significatives sont quelques peu différentes : nous observons la corrélation (positive) et significative entre l'effet du au traitement et la reconnaissance des macro-émotions de fierté, de joie (effet tendanciel), et entre les effets dus au traitement et la reconnaissance des micro-émotions (score total, score total des micros expressions négatives), et la reconnaissance des micro-émotions de peur de colère et de mépris. La reconnaissance des signaux véhiculés par la voix (score total voix) et par la voix (contenu RS) est positivement associée aux résultats.

La reconnaissance par les thérapeutes de la macro-émotion de soulagement, de la joie (micro) et du dégoût (micro) est associée significativement aux résultats du traitement en sens négatif.

Tableau 17. : Corrélations pour les scores CORE-10 ($n = 215$)

	Score de changement standardisé	
Reconnaissance Macro Fierté	$r = 0.14$	$p = 0.04$
Reconnaissance Macro Joie	$r = 0.14$	$p = 0.05$
Reconnaissance Macro Soulagement	$r = - 0.14$	$p = 0.04$
Reconnaissance des Micro expression (total)	$r = 0.16$	$p = 0.02$
Reconnaissance Micro émotions Négatives	$r = 0.21$	$p < 0.01$
Reconnaissance Micro-Joie	$r = - 0.18$	$p < 0.01$
Reconnaissance Micro-Peur	$r = 0.21$	$p < 0.01$
Reconnaissance Micro-Dégout	$r = - 0.20$	$p < 0.01$
Reconnaissance Micro-Colère	$r = 0.18$	$p < 0.01$
Reconnaissance Micro-Mépris	$r = 0.16$	$p = 0.02$
Reconnaissance Voix (total)	$r = 0.14$	$p = 0.04$
Reconnaissance Voix Contenu RS	$r = 0.20$	$p < 0.01$

Questionnaires PHQ-GAD combinés ($n = 171$, Tableau 18)

Nous retrouvons une association positive entre les scores de changement et la reconnaissance des signaux non verbaux suivants : la reconnaissance des macros-émotions, des macro-émotions positives, des macro-émotions négatives (colère, joie, peur, tristesse, irritation), des émotions négatives (micro), du dégoût (micro), des signaux corporels, et de la voix (RS). La capacité à décoder les signaux dans la voix (avec un contenu filtré) est, elle, paradoxalement associée négativement aux résultats du traitement.

Nous pouvons constater des différences dans les relation significatives entre le questionnaire CORE-10 qui mesure peut-être une dimension symptomatique plus générale (la

reconnaissance des macro-émotions positives semble être reliée à une meilleure réussite au traitement) et les questionnaires PHQ-9 et GAD-7 qui mesurent une dimension prétendument plus reliée à des affects négatifs, présents dans les troubles dépressifs et anxieux (au niveau macro ce sont les émotions négatives qui sont retrouvées comme prédictives).

Au niveau de la reconnaissance des micro-expressions, dans le cas du questionnaire CORE-10, toutes les émotions (positives et négatives) sont associées au changement alors que dans le cas de la combinaison PHQ-9 et GAD-7, seules les émotions négatives sont significatives.

Il est à noter que dans nos analyses, le score de flexibilité psychologique ou de pratique délibérée n'était pas relié significativement aux résultats pour le groupe total ($n=386$), cependant dans le groupe PHQGAD, la pratique délibérée était associée à de moins bons résultats ($r = -.18$; $p = .02$). Nous devons rappeler que les thérapeutes qui ont été évalués sur leur pratique délibérée lors du dernier mois, dans notre étude, n'ont pas rapporté que cette quantité de temps dévolue à leur pratique délibérée était habituelle, dans à peu près la moitié des cas, et que de ce fait, ces données ne sont pas représentatives de leur pratique au long cours ; ces résultats sont donc à prendre avec précaution.

Globalement, entre toutes les conditions, il apparaîtrait que la reconnaissance des émotions dans la voix (RS) soit associée positivement aux résultats de la psychothérapie.

Tableau 18. : Corrélation pour les scores PHQGAD ($n = 171$)

	Score de changement standardisé	
Reconnaissance Globale des macro-émotions	$r = 0.22$	$p < 0.01$
Reconnaissance des Macro émotions Positives	$r = 0.19$	$p = 0.02$
Reconnaissance des Macro émotions	$r = 0.23$	$p < 0.01$
Reconnaissance Macro Colère	$r = 0.20$	$p < 0.01$
Reconnaissance Macro Joie	$r = 0.18$	$p = 0.02$
Reconnaissance Macro Peur	$r = 0.16$	$p = 0.04$
Reconnaissance Macro Tristesse	$r = 0.19$	$p = 0.01$
Reconnaissance Macro Irritation	$r = 0.18$	$p = 0.02$
Reconnaissance Micro émotions Négatives	$r = 0.2$	$p < 0.01$
Reconnaissance Micro-Dégout	$r = 0.18$	$p = 0.02$
Reconnaissance Corps	$r = 0.22$	$p < 0.01$
Reconnaissance Voix (CF)	$r = - 0.20$	$p < 0.01$
Reconnaissance Voix (RS)	$r = 0.22$	$p < 0.01$

1.6.4.7. Calcul des corrélations entre les scores moyens des prédicteurs (associations entre la reconnaissance des différents signaux non verbaux et les scores moyens de changement ajustés par la modélisation multiniveau).

Afin d'obtenir des corrélations plus fiables et moins biaisées par l'effet de la répétition des données liées aux thérapeutes, nous avons calculé les corrélations au niveau des thérapeutes ($n = 11$), entre nos différents prédicteurs et le score de changement associé au thérapeute, et calculé par la modélisation multiniveau. Nous avons utilisé le test de corrélation en rang de Spearman du fait de l'échantillon réduit.

Sur l'échantillon total de départ, seul l'association entre la reconnaissance de la voix (RS) et les résultats était positivement significative ($\rho = .71, p = .02$), la corrélation entre les scores de reconnaissance des différents thérapeutes pour la détection de l'expression de la joie (macro) se rapprochant le plus, parmi nos autres prédicteurs, d'un effet tendanciel ($\rho = .59, p = .06$).

En considérant seulement la portion de l'échantillon initial de patients qui sont au-dessus du seuil clinique (score brut égal ou supérieur à 11 pour la CORE-10, score égal ou supérieur à 5 pour l'échelle PHQ-9 ou GAD-7), nous obtenons ainsi une population pour laquelle, l'indice de taille de l'effet qui pourra être calculé, est non seulement ajusté (par les résidus), mais également par le niveau de sévérité à l'entrée dans les soins, qui nous l'avons vu est un des prédicteur les plus importants de la réussite au traitement. Dans ce cas, la taille de l'effet calculée est appelée taille de l'effet ajustée, pour la sévérité qui est un indice qui a été inventé par le Dr Jebb Brown (Severity Adjusted Effect Size ou SAES, voir Minami et al., 2012). Sur l'échantillon restreint ainsi obtenu de 346 patients, l'association entre la reconnaissance de la voix (RS) et les résultats, était positivement significative, mais plus faible que sur l'échantillon total ($\rho = .62$; $p = .04$), et la corrélation entre les scores de reconnaissance des différents thérapeutes pour la détection de l'expression de la joie (macro) et les résultats au traitement, s'éloignait d'un effet tendanciel ($\rho = .56$; $p = .07$).

Enfin, en ajustant plus avant la sévérité clinique (avec un seuil plus restrictif à 13 points pour la CORE-10 et à 10 points pour la PHQ-9 ou pour la GAD-7), nous obtenons un échantillon final de 260 patients. Sur cet échantillon, l'association entre la reconnaissance de la voix (RS) et les résultats, n'était plus significative ($\rho = .40$, $p = .23$) et l'association entre la détection de l'expression de la joie (macro) et les résultats, s'éloignait encore plus sensiblement d'un effet significatif ($\rho = .42$; $p = .20$).

Nous pouvons évoquer que la diminution de la taille de l'échantillon puisse être susceptible de diminuer la puissance, c'est-à-dire la capacité à détecter un effet, s'il existe. S'y oppose toutefois le fait que la taille de l'effet du au traitement est attendue comme plus élevée dans une population dont le statut clinique serait plus sévère, et nos analyses de corrélation devraient donc permettre plus facilement de révéler des résultats significatifs, en restreignant notre échantillon à la frange exprimant un degré de détresse psychologique plus élevé.

Une corrélation n'est bien entendu pas l'équivalent d'un effet causal, mais nous pouvons émettre au final une autre hypothèse, en relation avec notre cadre théorique : l'hypothèse que la reconnaissance des signaux non verbaux puisse être significativement associée aux résultats, pour les patients les moins sévèrement déprimés, ou sub-syndromiques. En effet, pour eux, la manifestation de la détresse serait la moins marquée et accompagnée de moins de signaux non verbaux fiables de détresse. Les thérapeutes pourraient donc avoir plus de difficultés à jauger leur état ; ils seraient moins susceptibles de tirer des bénéfices réels des soins et ils pourraient plus bénéficier dans ces conditions, de l'adjonction d'un feedback instrumental indiquant au thérapeute qu'ils ne progressent pas. A l'opposé, un état clinique plus sévère, s'accompagnerait de signaux non verbaux plus saillants (notamment dans la « présentation du patient »), et l'effet de la relation entre la reconnaissance des signaux non verbaux et les résultats au traitement en serait fortement atténué.

1.6.4.8. Conclusions sur les hypothèses de départ.

Nous allons considérer seulement les résultats les plus conservateurs, c'est-à-dire non pas les analyses de corrélations exploratoires réalisées au niveau des patients, qui nous l'avons vu sont potentiellement biaisées, mais les résultats de nos dernières analyses entre les scores de changement au traitement ajustés par l'analyse multi-niveau, et les scores des capacités de nos thérapeutes à reconnaître les différents signaux non verbaux.

Dans cette étude, nous avons posé l'hypothèse principale que les habiletés des thérapeutes à décoder les signaux non verbaux puissent être significativement reliées aux résultats de la psychothérapie (H1). Les résultats vont dans le sens de nos hypothèses, en ce qui concerne la reconnaissance des signaux véhiculés par la voix uniquement (en condition RS-Réorganisation aléatoire des segments). L'effet de la joie est ici, tendanciel, et mériterait d'être exploré plus avant dans une nouvelle recherche.

Nos résultats préliminaires ne vont pas dans le sens des hypothèses H2, H3, et H4 puisque nous n'avons pas retrouvé de différence dans les relations entre la reconnaissance des différentes dimensions de signaux non verbaux, que ce soit au niveau des dimensions positives ou négatives de ces signaux, ou au niveau des macro-expressions et des micro-expression, en considérant seulement les données observées au niveau des thérapeutes.

Par contre, notre résultat préliminaire de lien entre la reconnaissance de signaux présents dans la voix (RS) et des résultats au traitement, va, nous le pensons, dans le sens de suggérer qu'il puisse exister un effet supplétif ou modérateur du feedback instrumental de la pratique ROM, aux signaux de feedback naturalistes présents dans les séances de psychothérapie délivrées par les thérapeutes de notre échantillon. En effet, tous les thérapeutes sont en condition de feedback instrumental, dans notre étude, mais l'effet d'une relation entre la reconnaissance des signaux non verbaux et les résultats au traitement est tout de même présent, et semble important (reconnaissance des signaux de la voix-RS). Un effet dû au feedback n'exclut pas que certains thérapeutes conservent une meilleure capacité discriminatoire des signaux non verbaux et encore une fois la corrélation n'est pas un effet causal. Toutefois la capacité de reconnaissance des signaux de la voix (RS) est liée aux résultats en condition feedback, l'effet feedback ne viendrait donc pas simplement annuler ou remplacer les signaux naturels et lisser les différences interindividuelles chez les thérapeutes dans la capacité à décoder les signaux non verbaux (ce qui correspondait à l'hypothèse palliative de l'effet du feedback instrumental ou au modèle GPS), mais viendrait plutôt, nous le postulons, interagir avec les capacités des thérapeutes à détecter la communication non verbale. De surcroît, l'association entre les résultats et la meilleure reconnaissance des signaux de la voix par les thérapeutes n'opèrerait que pour les patients qui sont les moins sévèrement affectés par la détresse psychologique, dans les mêmes conditions de traitement orienté par le feedback.

1.6.5. Discussion

Cette étude n'est bien entendu pas exempte de nombreuses limitations : d'abord les données ont été collectées de manière rétrospective et ce protocole de recherche n'autorise donc pas à tirer de conclusions de nature causale sur les « effets » (au sens statistique du terme) que nous aurions pu observer. Toutefois, nous pouvons noter que dans les recherches concernant les psychothérapies, les effets causaux peuvent assez difficilement être révélés dans les conditions naturelles du terrain, puisqu'ils nécessitent la variation systématique des conditions expérimentales, en toutes choses égales par ailleurs. Nous avons toutefois procédé à l'examen de relations entre des variables dont la mesure est censée être réputée temporairement stable (la reconnaissance des signaux non-verbaux), sans entraînement spécifique.

Ensuite nos échantillons sont de tailles réduites. Et bien que pour nos analyses initiales nous ayons veillé à ce qu'à chaque niveau d'analyse nous obtenions bien plus de dix unités, condition nécessaire à l'utilisation de ces modèles, nos analyses sur les échantillons pour lesquels la sévérité clinique avait été ajustée, ont pu pâtir de la diminution des effectifs pour chaque thérapeute.

Nos différents thérapeutes ont également administré des mesures de résultats différentes. Et bien que notre démarche de standardisation des effets soit statistiquement correcte, il est bien connu que des instruments de changement différents puissent produire des tailles d'effets, dues au même traitement, différentes. Avec des instruments mesurant des symptomatiques plus précises qui produiraient des tailles d'effets plus importantes. A ce sujet également, l'attrition importante entre le nombre de clients initialement disponibles et le nombre final de clients disponibles, dont les données complètes permettaient de calculer une taille de l'effet, ont pu être de nature à affecter nos résultats.

Nonobstant, cette étude est à notre connaissance la première étude visant à investiguer les relations entre les capacités à décoder les messages non-verbaux par les psychothérapeutes

et leurs résultats, pour des thérapeutes engagés dans une pratique ROM. De plus, il s'agissait d'utiliser une tâche de performance spécifiquement créée pour évaluer la capacité des thérapeutes à détecter les micro-expressions émotionnelles à l'aide de la présentation de « véritables micro-expressions » ce qui selon nous est également une démarche originale. Une extension de cette étude avec un design prospectif est d'ores-et-déjà considérée, avec la possibilité de fiabiliser la distribution de la tâche de reconnaissance des micro-expressions, et de la standardiser.

3. DISCUSSION GENERALE

Dans ce travail, nous avons initialement parcouru et synthétisé la littérature relative à l'efficacité de la psychothérapie qui apparaît réelle, forte et bien documentée. Il est également apparu que la part la plus importante dans la variance de résultats dus aux soins, était plutôt reliée aux thérapeutes et à leurs habiletés à interagir avec leurs patients, qu'à l'utilisation systématique d'un mode de traitement en particulier. Nous avons vu que les thérapeutes étaient exposés à des biais, et avaient tendance à surestimer leur efficacité, tout en présentant de grandes difficultés à détecter les patients à risque de ne pas s'améliorer, de se détériorer, voire d'abandonner le traitement. Ces thérapeutes ne disposaient pas en outre, de méthodes fiables pour développer leur performance qui atteindrait un plateau assez tôt au cours de leur carrière et accuserait même un déclin au cours de leurs carrière.

La méthodologie du Suivi Continu des Résultats de la psychothérapie (SCR- ROM) a été développée pour répondre à ces enjeux, en impliquant le patient dans sa thérapie, par la sollicitation de son feedback quant à son évolution symptomatique et quant à la qualité de la relation de soins. Cette pratique a initialement été documentée comme associée à des gains d'efficacité qui pouvaient être importants, et à des réductions des dépenses de santé. En analysant plus avant ces résultats, il est apparu que la pratique ROM n'apportait pas de bénéfices à tous les patients, que l'effet du feedback était lié à des variables modératrices reliées aux thérapeute (le genre, la propension au feedback, l'auto-efficacité), et que les mécanismes par lesquels le feedback instrumental affecterait les résultats des traitements psychothérapeutiques étaient méconnus. Le feedback instrumental a été initialement modélisé comme un signal (négatif) aidant les thérapeutes à infléchir la course des soins pour atteindre les buts fixés en commun avec leurs patients. Le signal de feedback agissant à la manière d'un système de navigation GPS et permettant de reprendre le chemin d'une thérapie efficace. En ce sens, il était logique que l'adjonction du feedback instrumental puisse affecter les résultats des

psychothérapies pour lesquelles la progression n'était initialement pas satisfaisante (pour les patients NOT – « Pas en bonne voie de rétablissement »).

Pourtant, tous les thérapeutes ne bénéficiaient pas de l'adjonction du feedback instrumental, et les thérapeutes initialement les moins efficaces pouvaient, dans certaines études, atteindre les résultats de leurs collègues naturellement plus efficaces, ce qui était de nature à suggérer que le feedback instrumental pourrait venir pallier ou compléter les habiletés des thérapeutes. Pour Worthen & Lambert (2007), le feedback ROM pourrait ainsi agir en apportant une information à laquelle les thérapeutes auraient initialement prêté peu d'attention.

D'où notre volonté de mieux comprendre quelle pouvait être cette source d'information ou dimension « cachée » pour laquelle certains thérapeutes pourraient avoir des capacités de perception différentes. En effet, la psychothérapie étant avant tout une entreprise relationnelle opérant à partir d'une communication d'influence, et à partir d'informations véhiculées par des messages verbaux et paraverbaux, qui sont en interaction permanente, notre analyse de la littérature nous a amené à supposer que le feedback instrumental de la pratique ROM pouvait représenter un remplaçant (hypothèse palliative) ou un complément (hypothèse supplétive) aux messages échangés au cours des séances et plus particulièrement aux messages véhiculés par des médias non-verbaux ; messages pour lesquels les thérapeutes pouvaient avoir des capacités de perceptions variables.

Notre objectif principal (étude numéro 1) était ainsi de procéder à l'adaptation française du système ROM PCOMS (échelles ORS et SRS) afin de permettre à la communauté francophone d'adopter une pratique délibérée du feedback, de documenter les qualités psychométriques de ce système et de fournir des données normatives préliminaires. En espérant faciliter de nouvelles recherches utilisant ce système.

Notre second objectif était de commencer à soumettre à l'épreuve des faits ou plutôt ici des données, nos hypothèses sur les relations qu'entretiendrait le feedback instrumental et les autres feedback naturalistes (non verbaux) qui sont rencontrés dans les séances de

psychothérapie, en testant les habiletés à reconnaître les signaux non verbaux de psychothérapeutes qui sont engagés dans une pratique ROM et qui enregistrent donc leur efficacité (étude numéro 2).

Dans l'étude numéro 1, nous avons ainsi réalisé la première adaptation française d'un système ultra-bref de Routine Outcome Monitoring (ROM) enregistré au répertoire américain de la SAMSHA. Les échelles ORS et SRS adaptées en français, ont nous l'avons vu, des qualités psychométriques comparables aux échelles américaines, avec une structure factorielle et une validité convergente comparable, des coefficients de fidélité très proches, et une sensibilité au changement satisfaisante qui permet, pour l'échelle ORS, de distinguer les personnes qui appartiennent à la population clinique, des personnes qui ne suivent pas un traitement. La progression des patients au cours des soins peut également être suivie à l'aide de l'échelle ORS. L'échelle SRS paraît également adaptée pour effectuer le suivi du rapport collaboratif en cours de traitement. Les données normatives préliminaires et les indices de changement fiables vont dans le sens des autres études de validation étrangères (Janse et al., 2014, Moggia et al., 2018) et cadrent avec les variations qui sont attendues lors d'une validation transculturelle.

Il est pourtant à noter que la solution factorielle obtenue avec les scores des items de l'échelle SRS est plus satisfaisante que celle que nous avons obtenu avec les scores des items de l'échelle ORS. En effet, nous avons rencontré un « Heywood Case » dans l'analyse factorielle confirmatoire de cette dernière. En étudiant de plus près les résultats, et en rapport avec la littérature documentant les qualités psychométriques de cette échelle en pratique de soins (Campbell & Hemsley, 2009), nous avons observé qu'il existait une situation de multi colinéarité entre les scores obtenus aux items a (Bien-être personnel) et d (Globalement -Bien-être général) de cette échelle. Cette multicolinéarité étant de nature à rendre la solution factorielle unidimensionnelle moins satisfaisante. Bien que cette situation ait été résolue par l'analyse statistique (avec une contrainte sur l'item d), nous pensons, en relation avec notre pratique clinique, qu'il pourrait être intéressant dans une perspective d'optimisation de la

mesure, de considérer les qualités psychométriques d'une échelle de résultats ne comprenant que les trois premiers items (a, b, c) de l'ORS, dans de nouvelles études.

En outre, bien que la méthode de l'analyse des corrélations entre les scores de deux échelles censées mesurer le même construit soit une méthode communément admise, nous pensons comme certains de nos confrères que cette méthode n'est pas la plus fiable (Seidel et al., 2017). En effet, documenter que les scores de deux instruments de mesure varient dans le même sens, ne revient pas à dire que ces deux instruments mesurent le même construit. La taille et le poids des individus varient par exemple dans le même sens, sur une population, parce qu'elles rendent compte en partie d'un autre construit qui leur est relié (la croissance). Nous avons pensé adopter une autre méthode, la méthode de Bland Altman qui permet de comparer deux techniques de mesure d'une même grandeur (Grenier, Dubreuil & Journois, 2000) ; et peut donc être utilisées pour analyser la convergence des scores de deux questionnaires, à condition que ces questionnaires mesurent effectivement la même grandeur. Dans le cas du questionnaire ORS par exemple et de l'échelle MI-22, les intervalles de score étaient différents. Intuitivement, les scores bruts auraient alors pu être transformés en une métrique commune (en scores standardisés Z en utilisant la moyenne et l'écart-type de nos échantillons), mais dans ce cas les scores auraient été centrés et réduits et l'étape de test de la différence des moyennes de la procédure de Bland-Altman n'aurait pas pu être réalisée.

Nous avons rencontré d'autres difficultés et des limitations dans notre démarche : d'abord, malgré une campagne de recrutement active et importante, avec de nombreux appels à la participation à notre recherche auprès de groupement de thérapeutes, d'organisations de psychologues, et de nombreux mailing à des psychologues contactés directement en France et en Suisse, nous avons eu assez peu de praticiens qui se sont manifestés pour prendre part à l'étude numéro 1. Sept thérapeutes devaient initialement prendre part à notre protocole et trois thérapeutes n'ont finalement pas participé (pour des raisons de patientèle ou de population suivies au final différentes de nos critères d'inclusion, et pour d'autres raisons non-

documentées). Il apparaît évidemment difficile de « partager » les résultats de sa pratique clinique, même dans le cas de recherches où l'anonymat est garanti ; et même en visant la réduction du temps « englouti » par ce protocole, et en offrant une formation à la pratique ROM aux praticiens à la fin de l'étude, il fut difficile de recruter des cliniciens.

Nous avons également rencontré des difficultés techniques, puisque nous avons eu des problèmes dans l'administration des questionnaires depuis le serveur de l'université (par exemple une panne informatique lors d'un épisode caniculaire). Le logiciel Limesurvey a pu comme toute solution informatique, poser des difficultés techniques (blocage de l'interface chez le thérapeute, enregistrement vides, problèmes dans l'enregistrement du numéro de la séance) qui nous a amené à assurer un support technique auprès des thérapeutes et des étudiants, et nous a amené également à consolider les fichiers de données finalement analysés.

Toutefois, nous avons également dépassé certaines limitations qui étaient liées aux recherches antérieures sur les qualités psychométriques des échelles PCOMS. Nous avons par exemple documenté les tailles de l'effet pour nos analyses, en plus de la significativité.

Nous avons produit des analyses factorielles confirmatoires, les structures factorielles de ces échelles étant peu souvent documentées dans les autres recherches. Nous avons également produit des comparaisons entre les différents échantillons (cliniques en début à l'entrée au traitement, clinique en cours de traitement, et non cliniques), en condition feedback et non feedback.

Nous avons également évalué la validité critérielle de l'échelle ORS en relation avec le trait d'affectivité négative mesuré par les Big-Five (l'affectivité négative ou détresse trait) qui est reliée à l'état de détresse psychologique, ce qui à notre connaissance, n'avait pas été encore réalisé.

Nous avons également utilisé la méthodologie ROC pour compléter nos données sur le seuil de l'échelle ORS. Enfin, nous avons voulu nous rapprocher au mieux des

recommandations de la commission internationale des tests (ITC) en standardisant l'administration des questionnaires à l'aide du support informatique, et nous avons également comparé les conditions de renvoi ou non du feedback en « aveuglant » (blind) le thérapeute face au feedback du patient. Nous avons également utilisé des méthodes statistiques dites robustes qui sont les plus adaptées pour traiter des échantillons de données réduits.

Au-delà de ces considérations, nous nous interrogeons encore face à ces données : pourquoi les corrélations sont-elles plus élevées entre les scores obtenus aux items de l'échelle ORS et les scores obtenus aux échelles de référence pour la détresse psychologique (MI-22) dans l'échantillon non clinique ? Dans cet échantillon, la détresse psychologique est moins élevée, l'ORS discriminerait mieux de ce fait la détresse psychologique quand elle est moins élevée, et cette échelle pourrait au final, ce qui est en concordance avec son utilisation, mesurer un degré de bien-être et son augmentation au cours de la thérapie. Le degré de bien-être n'étant pas le pôle positif d'un construit pour lequel le pôle négatif serait le degré de mal être (détresse psychologique).

Enfin, nous avons remarqué que la consistance interne calculée sur les scores des items de la SRS augmentait en cours de traitement (Hafkenscheid 2010, Janse et al., 2014). Nous pouvons nous interroger sur la signification de ce résultat et supposer que le développement du rapport collaboratif ne se réalise pas de manière discrète et dichotomisée (sur une dimension), bien que l'analyse statistique puisse permettre de traiter les données des items séparément. Quand le score augmenterait sur une dimension liée à l'alliance, la qualité globale du rapport collaboratif serait affectée, ce qui mériterait selon nous d'être exploré.

Dans l'étude numéro 2, nous avons envisagé des liens entre la capacité des psychothérapeutes engagés dans une pratique ROM à reconnaître des signaux de communication non verbale et les résultats obtenus auprès de leurs patients. Dans ce travail exploratoire, nous avons pu mettre à l'épreuve des faits, l'association possible entre une meilleure reconnaissance des signaux non verbaux, pour des modalités de présentation

différentes (macro-expressions émotionnelles, micro-expressions émotionnelles, signaux véhiculés par la voix, par le corps, et par les gestes).

Ce travail n'est pas non plus exempt de limitations. D'abord, l'échantillon de thérapeutes sondés dans cette étude était restreint et la population de patients était également réduite. Notre protocole ne nous a pas permis d'interpréter de relations dynamiques (c'est-à-dire se rapprochant d'une quelconque interprétation causale) puisqu'il s'agissait d'une étude rétrospective concernant des patients dont les données avaient été archivées. Il s'agissait essentiellement d'explorer les corrélations entre nos variables d'intérêt en ayant au préalable écarté les variables qui dans la littérature n'étaient pas considérées comme des variables d'influence du phénomène que nous avons étudié. Ceci n'exclut pas que ces relations ne puissent être influencées par un « tertium quid », c'est-à-dire une variable fantôme, qui influencerait à la fois les résultats thérapeutiques, mais également les capacités des thérapeutes à détecter les signaux non-verbaux (nous pouvons évoquer le rôle vraisemblablement envisageable de la fatigue, de la concentration, et de l'humeur du thérapeute).

Cependant, nous nous sommes assurés que les construits que nous avons opérationnalisés reflétaient des mesures stables, sans entraînement formel, et que nos thérapeutes n'avaient pas été formés, en outre, à développer ces compétences, ce qui aurait pu brouiller la signification de nos résultats.

Notre résultat principal d'une association, retrouvée chez nos onze thérapeutes, entre la reconnaissance des signaux de la voix (avec une réorganisation temporelle aléatoire des segments) et les résultats au traitement pourrait être due à un artefact statistique (la similarité des praticiens à détecter ces signaux). Et ce résultat pourrait ne pas être retrouvé dans une population plus importante de thérapeutes.

Pourtant nous pouvons quand même nous interroger sur la nature des informations qui pourraient être véhiculées par les enregistrements vocaux qui auraient subi un tel traitement.

Nous n'avons pas en effet jusqu'ici, développé comment ce traitement de l'information (randomized splicing) altérerait les qualités des signaux de la voix. Dans son ouvrage « Voice in social interaction », Pittam (p115) explique que cette technique, développée par Scherer en 1971, s'inscrit dans les stratégies visant à neutraliser le contenu sémantique de la voix. Cette stratégie est ainsi supposée altérer les indices temporels des signaux présents dans la voix, sans altérer la qualité de la voix, son amplitude et les fréquences fondamentales ; avec une efficacité établie, face à d'autres techniques de filtrage du contenu, permettant notamment la reconnaissance de la colère, de la tristesse et de la peur, après filtrage.

Comment un tel traitement de la voix qui éliminerait des indices temporels permettrait-il à des êtres humains de distinguer des signaux non verbaux ? Cette question impliquerait selon nous de nouvelles recherches. Nous postulerons simplement que si cet effet d'association entre les résultats de la psychothérapie et de reconnaissance de la voix (RS) existe véritablement, pour les patients dont la détresse psychologique n'est pas fortement marquée, il pourrait être en lien avec la sensibilité des cliniciens à détecter l'incongruence entre les messages verbaux et non verbaux (de la voix) ; et à répondre fonctionnellement aux patients. Une approche en analyse des modérateurs et des médiateurs, manipulés lors d'un entraînement, et donc avec une séquence prospective, serait à même selon nous, d'apporter de nouveaux éléments de réponse à ces questions.

Il est déjà possible d'utiliser le système PCOMS et la pratique ROM pour aider à encadrer la délivrance des soins et à fournir des données probantes aux organismes d'assurance-santé en permettant de documenter l'efficacité de soins (en vue d'une prise en charge des coûts), et pour rediriger les patients vers d'autres professionnels de la santé en cas de mauvaise réussite initiale au traitement dans un continuum de soins. Et finalement pour aider à développer son efficacité clinique (cliniciens, institutions).

Les informations issues de la pratique ROM pourraient en ce sens apporter des données probantes quant à l'efficacité des psychothérapies en pratique clinique de routine, en vue

d'encadrer un futur remboursement des soins par les organismes d'assurance-santé, qui est au cœur des débats actuels. Les informations à transmettre pouvant se limiter aux effets du traitement et exclure des données plus personnelles sur le patient. Ces données pourraient également aider les cliniciens à prendre des décisions cliniques, au cours de la délivrance des traitements, en termes de fréquence des séances, de pause thérapeutique, d'arrêt du traitement et de réorientation des patients vers des confrères ou des institutions.

Une meilleure compréhension des effets dus au feedback en psychothérapie est susceptible de déboucher, selon nous, sur l'amélioration des modes de délivrances des soins. Les prolongements de ce travail exploratoire sont d'ores et déjà considérés.

Nous souhaitons en conclusion, renvoyer le lecteur à la phrase attribuée à Lao-tseu et évoquée au début de cette thèse : Que savons-nous vraiment, et quels moyens avons-nous de le savoir ? En effet, de multiples « découvertes scientifiques » autrefois considérées comme acquises ont depuis été réfutées (par exemple l'impossibilité de la neurogènes qui fait pourtant encore débat¹⁵). Et même si pour citer Heraclite « Tout Coule », se donner les moyens de mieux comprendre, reste quand même, selon nous, une entreprise extrêmement importante.

¹⁵ <https://lejournald.cnrs.fr/nos-blogs/aux-frontieres-du-cerveau/le-cerveau-humain-adulte-ne-produirait-pas-de-nouveaux-neurones>

4. REFERENCES BIBLIOGRAPHIQUES

- Alexander F. et French T.M. (1946), *Psychoanalytic therapy: principles and application*, N.Y. Ronald Press
- Altmann, U., Thielemann, D., Zimmermann, A., Steffanowski, A., Bruckmeier, E., Pfaffinger, I., Strauß, B. (2018). Outpatient Psychotherapy Improves Symptoms and Reduces Health Care Costs in Regularly and Prematurely Terminated Therapies. *Frontiers in Psychology*, 9, 748. <http://doi.org/10.3389/fpsyg.2018.00748>.
- Anderson, T., Crowley, M. E. J., Himawan, L., Holmberg, J. K., & Uhlin, B. D. (2015). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research*, 26, 5, 511-529.
- Anderson, T., Knobloch-Fedders, L. M., Stiles, W. B., Ordonez, T., & Heckman, B. D. (2012). The power of subtle interpersonal hostility in psychodynamic psychotherapy: A speech acts analysis. *Psychotherapy Research*, 22, 3, 348-362.
- Anderson, E. M., & Lambert, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology*, 57, 7, 875-888.
- Anderson, T., Ogles, B.M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist Effects: Facilitative Interpersonal Skills as a Predictor of Therapist Success. *Journal of Clinical Psychology*, 65, 755-768.
- APA Presidential Task Force on Evidence-Based Practice. (2006). *Evidence-based practice in psychology*. *American Psychologist*, 61(4), 271–285.
- Arnaud, B., Malet, L., Teissedre, F., Izaute, M., Moustafa, F., Geneste, J., Schmidt, J., ... Brousse, G. (2010). Validity Study of Kessler's Psychological Distress Scales Conducted Among Patients Admitted to French Emergency Department for Alcohol Consumption-Related Disorders. *Alcoholism: Clinical and Experimental Research*, 34, 7, 1235–1245.

- Asay, T. R., & Lambert, M. J. (1999). The empirical case of the common factors in psychotherapy: quantitative findings. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: what works in therapy* (pp. 23-55). Washington DC: American Psychological Association. doi:10.1037/11132-001.
- Ashaye OA, Livingston G, Orrell MW. Does standardized needs assessment improve the outcome of psychiatric day hospital care for older people? A randomized controlled trial. *Aging Ment Health* 2003; 7: 195–9.
- Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). *Using client feedback to improve couple therapy outcomes: a randomized clinical trial in a naturalistic setting. Journal of Consulting and Clinical Psychology, 77, 4, 693-704.*
- Bachelor, A., & Horvath, A. (1999). The therapeutic relationship. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 133–178). Washington, DC: APA Press.
- Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the Alliance-Outcome Correlation: Exploring the Relative Importance of Therapist and Patient Variability in the Alliance. *Journal of Consulting and Clinical Psychology, 75, 6, 842–852.*
- Bauer S. Outcome Monitoring and Feedback: Strategies for Optimisation of Inpatient Psychotherapy? (in German). Doctoral Thesis (PhD), Faculty of Social and Behavioural Sciences, University of Tübingen, Germany, 2004.
- Bänziger, T., Scherer, K.R., Hall, J.A., Rosenthal, R. (2011). Introducing the MiniPONS: A Short Multichannel Version of the Profile of Nonverbal Sensitivity (PONS). *Journal of Nonverbal Behavior, 35, 189–204.*
- Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., ... Evans, C. (2013). *The CORE - 10: A short measure of psychological distress for routine use in the psychological therapies.* (Counselling and psychotherapy research)
- Barkham, M., & Margison, F. (2008). Practice-based Evidence as a Complement to Evidence-based Practice: From Dichotomy to Chiasmus. 443-476.

- Baudrant-Boga, M., Lehmann, A., & Allenet, B. (2012). Penser autrement l'observance médicamenteuse : d'une posture injonctive à une alliance thérapeutique entre le patient et le soignant – Concepts et déterminants. *Annales Pharmaceutiques Françaises*, 70(1), 15–25. doi:10.1016/j.pharma.2011.10.003.
- Berking M, Orth U, Lutz W. Effects of systematic feedback to the therapist on patient progress. An empirical study in a cognitive-behavioral inpatient setting (in German). *Psych Psychot* 2006; 35: 21–9.
- Benish, S., Imel, Z. E., & Wampold, B. E. (2008). The relative efficacy of bona fide psychotherapies of post-traumatic stress disorder: *A meta-analysis of direct comparisons*. *Clinical Psychology Review*, 28, 746-758.
- Bertolino, B., Bargmann, S., & Miller, S. (2012). International Center for Clinical Excellence manuals on feedback-informed treatment (FIT): Manual 1--What works in therapy: A primer. Chicago, IL: International Center for Clinical Excellence.
- Bickman, L., Kelley, S. D., & Athay, M. (2012). The Technology of Measurement Feedback Systems. *Couple & Family Psychology*, 1, 4, 274–284.
- Bickman, L., Kelley, S. D., Breda, C., de, A. A. R., & Riemer, M. (2011). *Effects of Routine Feedback to Clinicians on Mental Health Outcomes of Youths: Results of a Randomized Trial*. *Psychiatric Services*, 62, 12, 1423.
- Bond, F. W., Lloyd, J. and Guenole, N. (2013), The work-related acceptance and action questionnaire: Initial psychometric findings and their implications for measuring psychological flexibility in specific contexts. *J Occup Organ Psychol*, 86: 331-347. doi:10.1111/joop.12001.
- Bordin, E. (1979). *The generalizability of the psychoanalytic concept of the working alliance*. *Psychotherapy: Theory, Research and Practice*, 16, 252-260.
- Bringhurst, D. L., Watson, C. S., Miller, S. D., & Duncan, B. L. (2006). The reliability and validity of the outcome rating scale: A replication study of a brief clinical measure. *Journal of Brief Therapy*, 5 (1), 23–29.

- Brodey BB, Cuffel B, McCulloch J, Tani S, Maruish M, Brodey I, et al. The acceptability and effectiveness of patient-reported assessments and feedback in a managed behavioral healthcare setting. *Am J Manag Care* 2005; 11: 774–80.
- Brown, J. Dreis, S., & Nace, D. (1999). What really makes a difference in psychotherapy outcome? Why does managed care want to know? In M. Hubble, B. Duncan, & S. Miller (Eds.), *The heart and soul of change* (pp. 389–406). Washington, DC: APA Press.
- Brown, G. S., & Lambert, M. J. (1998). *Tracking patient progress: Decision making for cases who are not benefiting from psychotherapy*. Paper presented at the annual meeting of the Society for Psychotherapy Research, Snowbird, UT.
- Campbell, A., & Hemsley, S. (2009). Outcome Rating Scale and Session Rating Scale in psychological practice : Clinical utility of ultra-brief measures. *Clinical Psychologist*, 13, 1, 1–9.
- Campbell, L. F., Norcross, J. C., Vasquez, M. J., & Kaslow, N. J. (2013). *Recognition of psychotherapy effectiveness: the APA resolution*. *Psychotherapy (chicago, Ill.)*, 50, 1, 98-101.
- Carlier, P. D., Meuldijk, P. D., Van, V. M., Van, F. M., Van, . W. M., & Zitman, M. D. (2012). *Routine outcome monitoring and feedback on physical or mental health status: evidence and theory*. *Journal of Evaluation in Clinical Practice*, 18, 1.)
- Carr, A., & Irish Council for Psychotherapy. (2007). *The effectiveness of psychotherapy: A review of research*. Dublin: Irish Council for Psychotherapy.
- Carver, C. S., & Scheier, M. F. (1981). *Attention and Self-Regulation: A Control Theory Approach to Human Behavior*. New York: Springer. <http://dx.doi.org/10.1007/978-1-4612-5887-2>.
- Carver, C. S., & Scheier, M. F. (1982). *Control theory: A useful conceptual framework for personality-social, clinical, and health psychology*. *Psychological Bulletin*, 92, 1, 111-135.

- Caspi, O. (2004). How good are we? A meta-analytic study of effect sizes in medicine. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 65(5-B), 2607.
- Cautin, R. L. A century of psychotherapy, 1860-1960. In Norcross, J. C., VandenBos, G. R., Freedheim, D. K., & American Psychological Association. (2011). *History of psychotherapy: Continuity and change*. Washington, DC: American Psychological Association.
- Cazauvieilh, C. (2013, novembre). *Évaluer et améliorer l'efficacité des psychothérapies : adaptation française d'un système de suivi des résultats thérapeutiques et étude prospective des processus participants au changement*. Communication publique présentée au séminaire de l'école doctorale SP2, Université Bordeaux II, Bordeaux.
- Cazauvieilh, C. *Améliorer son efficacité clinique et dynamiser sa croissance professionnelle : des méthodes simples au service du changement*. Session de poster, vendredi 12 décembre 2014, congrès de l'Association Française de Thérapie Comportementale et Cognitive (AFTCC), Paris, maison de la chimie.
- Cazauvieilh, C. *Fonctionnement et dysfonctionnement des psychothérapies*. Session Clinique, le vendredi 4 Septembre 2015. 56^{ème} Congrès de la Société Française de Psychologie. Université de Strasbourg.
- Cazauvieilh, C (2015). De la lutte pour ressentir à l'acceptation pour mieux agir. Une approche ACT intégrative du traitement des attaques de panique et des expériences de dépersonnalisation-déréalisation. In Seznec, J.C. (Ed.), *ACT : applications thérapeutiques*. Paris: DUNOD.
- Chang TE, Jing Y, Yeung AS, Brennenman SK, Kalsekar I, Hebden T, et al. Effect of communicating depression severity on physician prescribing patterns: findings from the Clinical Outcomes in MEasurement-based Treatment (COMET) trial. *General Hospital Psychiatry* 2012;34:105-12. [DOI: 10.1016/j.genhosppsy.2011.12.003].
- Chiles, J., Lambert, M. J., & Hatch, A. L. (1999). The impact of psychological interventions on medical cost offset: A meta-analytic review. *Clinical Psychology*, 6(2), 204–220.

- Chow, D. L. (2014). *The study of supershrinks: Development and deliberate practices of highly effective psychotherapists*. Doctoral Thesis presented for the degree of Doctor of Philosophy of the Curtin University
- Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy, 52*, 3, 337-345.
- Claiborn, C. D., & Goodyear, R. K. (2005). *Feedback in psychotherapy*. *Journal of Clinical Psychology, 61*, 2, 209-17.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conrath, P. & Houdry, P. (2017). Les psychologues, remboursés ? *Le Journal des psychologues, 343*,(1), 3-3. doi:10.3917/jdp.343.0003.
- Consoli, S. M. (2006). Alliance thérapeutique. *Annales d'Endocrinologie, 67*(6), 645–647. doi:10.1016/s0003-4266(06)78350-0.
- Cooper, M. (2008). *Essential Research Findings in Counselling and Psychotherapy*. London: Sage
- Corbière, M., Bisson, J., Lauzon, S., & Ricard, N. (2006). Factorial validation of a French short-form of the Working Alliance Inventory. *International Journal of Methods in Psychiatric Research, 15*, 1, 36–45.
- Crawford, M., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., Parry, G. (2016). Patient experience of negative effects of psychological treatment: Results of a national survey. *British Journal of Psychiatry, 208*(3), 260-265. doi:10.1192/bjp.bp.114.162628.
- Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A Meta-Analysis of Cognitive-Behavioural Therapy for Adult Depression, Alone and in Comparison With Other Treatments. *Canadian Journal of Psychiatry, 58*, .)

- Davidson, A. H., Poulsen, S., Lindschou, J., Winkel, P., Tróndarson, M. F., Waaddegaard, M., & Lau, M. (2017). Feedback in group psychotherapy for eating disorders: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 85*(5), 484-494. <http://doi.org/10.1037/ccp0000173>.
- Davison A.; C. Lansley; N. Costen; K. Tan; M. H. Yap, "SAMM: A Spontaneous Micro-Facial Movement Dataset," in IEEE Transactions on Affective Computing, vol.PP, no.99, pp.1-1 doi: 10.1109/TAFFC.2016.2573832
- Ekman, P., W.V. Friesen, and Joseph C. Hager. Facial Action Coding System: The Manual. Facial Action Coding System, user & investigator guides. A Human Face, 2002.
- Davidson, K., Perry, A., & Bell, L. (2015). Would continuous feedback of patient's clinical outcomes to practitioners improve NHS psychological therapy services? Critical analysis and assessment of quality of existing studies. *Psychology and Psychotherapy, 88*(1), 21-37. <http://doi.org/10.1111/papt.12032>.
- Dazzi T, Gribble R, Wessely S, Fear NT. Does asking about suicide and related behaviours induce suicidal ideation? What is the evidence? *Psychol Med. 2014;44:3361–3*. [PubMed].
- Dezetter, A., Briffault, X., Ben Lakhdar, C., & Kovess-Masfety, V. (2014). Costs and Benefits of Improving Access to Psychotherapies for Common Mental Disorders. *The Journal of Mental Health Policy and Economics, 16*, 161–177.
- Duncan, B. L., Miller, S. D., Wampold, B. E., & Hubble, M. A. (Eds.). (2010). *The heart and soul of change: Delivering what works in therapy* (2nd ed.). Washington, DC: American Psychological Association.
- De Jong, K., van, S. P., Nugter, M. A., Heiser, W. J., & Spinhoven, P. (2012). *Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. Psychotherapy Research, 22*, 4, 464-474.
- Dowrick C, Leydon GM, McBride A, Howe A, Burgess H, Clarke P, et al. (2009) Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *BMJ;338:b663*. doi:10.1136/bmj.b663.

Ekman, P. (2003). *Emotions revealed* (2nd ed.). New York: Times Books.

Ekman, P. and Friesen, W.V. (1969) The Repertoire or Nonverbal Behavior: Categories, Origins, Usage and Coding. *Semiotica*, 1, 49-98.
<https://doi.org/10.1515/semi.1969.1.1.49>.

Ekman, P. & Friesen, W. V. (1974). Nonverbal Behavior and Psychopathology. In Friedman, R. J. & Katz, M. M. (Eds.), *The Psychology of Depression: Contemporary Theory and Research* (pp. 203-232). Washington, D.C.: Winston & Sons.

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A Few Can Catch a Liar. *Psychological Science Cambridge-*, 10, 3, 263-266.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., ... et, (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Archives of General Psychiatry*, 46, 11, 971-82.

Ericsson, K. A., Charness, N., Feltovich, P., & Hoffman, R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.

Eysenck, H. J. (1952). The effects of psychotherapy: an evaluation. *Journal of Consulting Psychology*, 16(5), 319-324.

Flynn, R.J., Aubry, T.D., Guindon, S., Tardif, L., Viau, M. Et Gallant, A. (2003). Validation d'une version française abrégée du Outcome Questionnaire et évaluation d'un service de counselling en milieu clinique, *Canadian Journal of Program Evaluation*, 17(3), 57-74.

Frank, J. D., & Frank, J. B. (1993). *Persuasion and healing: A comparative study of psychotherapy*. Baltimore: Johns Hopkins University Press.

International Test Commission (2014). International Guidelines on the Security of Tests, Examinations, and Other Assessments. [www.intestcom.org] -French.

- Friedman, H. S. (1978). The Relative Strength of Verbal versus Nonverbal Cues. *Personality and Social Psychology Bulletin*, 4, 1, 147-150.
- Gerst-Emerson, K., & Jayawardhana, J. (2015). Loneliness as a Public Health Issue: The Impact of Loneliness on Health Care Utilization Among Older Adults. *American Journal of Public Health*, 105(5), 1013–1019. <http://doi.org/10.2105/AJPH.2014.302427>.
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74-78.
- Goldberg, S. B., Babins-Wagner, R., Rousmaniere, T., Berzins, S., Hoyt, W. T., Whipple, J. L., Miller, S. D., ... Wampold, B. E. (2016). Creating a climate for therapist improvement: A case study of an agency focused on outcomes and deliberate practice. *Psychotherapy*, 53, 3, 367-375.
- Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E. (2016). Do Psychotherapists Improve With Time and Experience? A Longitudinal Analysis of Outcomes in a Clinical Setting. *Journal of Counseling Psychology*, 63, 1, 1-11.
- Grenon, R., Carlucci, S., Brugnera, A., Schwartze, D., Hammond, N., Ivanova, I., Mcquaid, N., Tasca, G. A. (2018). Psychotherapy for eating disorders: A meta-analysis of direct comparisons. *Psychotherapy Research : Journal of the Society for Psychotherapy Research*, 2018, 1-13.
- Hafkenschied, Anton & Duncan, Barry & Miller, Scott. (2010). The Outcome and Session Rating Scales: A Cross-Cultural Examination of the Psychometric Properties of the Dutch Translation. *Journal of Brief Therapy*. 7 (1 & 2).
- Haggard E.A., Isaacs K.S. (1966) Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy. In: *Methods of Research in Psychotherapy*. The Century Psychology Series. Springer, Boston, MA

- Hall, E. T. (1969). *The hidden dimension*. Garden City, N.Y.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85(4), 845-857. <http://dx.doi.org/10.1037/0033-2909.85.4.845>
- Hall, J. A., & Bernieri, F. J. (2001). *Interpersonal sensitivity: Theory and measurement*. Mahwah, NJ: Erlbaum.
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *J. Clin. Psychol.*, 61(2), 155–163. doi:10.1002/jclp.20108
- Harmon, C., Lambert, M., Smart, D., Nielsen, Slade, K., Lutz, W. (2007). *Psychotherapy Research*, 17(4), 379-392.
- Hatfield, D. , McCullough, L. , Frantz, S. H. and Krieger, K. (2010), *Do we know when our clients get worse? an investigation of therapists' ability to detect negative client change*. *Clin. Psychol. Psychother.*, 17: 25-32. doi:[10.1002/cpp.656](https://doi.org/10.1002/cpp.656)
- Hawkins EJ, Lambert MJ, Vermeersch DA, Slade KL, Tuttle KC. The therapeutic effects of providing patient progress information to therapists and patients. *Psychother Res* 2004; 14: 308–27.
- Hayes S., Pistorello J., Levin M., (2012). Acceptance and Commitment Therapy as a Unified Model of Behavior Change. *The Counseling Psychologist*. 40. 976-1002. 10.1177/0011000012460836.
- Herold, D. M., & Fedor, D. B. (2003). *Individual differences in feedback propensities and training performance*. *Human Resource Management Review*, 13, 4, 675-689.
- Hill, C. E., Anderson, T., Kline, K., McClintock, A., Cranston, S., McCarrick, S., Petrarca, A., ... Gregor, M. (2016). Helping Skills Training for Undergraduate Students: Who Should We Select and Train?. *The Counseling Psychologist*, 44, 1, 50-77.

- Hill, C., Thompson, B., & Corbett, M. (1992). The Impact of Therapist Ability to Perceive Displayed and Hidden Client Reactions on Immediate Outcome in First Sessions of Brief Therapy. *Psychotherapy Research*, 2, 2, 143-155.
- Holt-Lunstad, J., Baker, M., Harris, T., Stephenson, D., & Smith, T. B. (2015). Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review. *Perspectives on Psychological Science*, 10, 2, 227-237.
- Horvath, A. O., & Bedi, R. P. (2002). The Alliance. In J. Norcross (Ed.), *Psychotherapy Relationships That Work: Therapist Contributions and Responsiveness to Patients* (pp. 37-70). New York: Oxford University Press. Psychotherapy. Oxford, New York (Reviewed).
- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, 41(2), 159–164. doi:10.1037/0003-066X.41.2.159.
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy. Efficacy, effectiveness, and patient progress. *The American Psychologist*, 51, 10, 1059-64.
- Hwang, H. S., & Matsumoto, D. (2010). *Training improves the ability to recognize subtle facial expressions of emotion*. Paper presented at the Annual Convention of the Association for Psychological Science, Boston, MA.
- Imel, Z. E., Wampold, B. E., Miller, S. D., & Fleming, R. R. (2008). Distinctions without a difference: Direct comparisons of psychotherapies for alcohol use disorders. *Journal of Addictive Behaviors*, 22, 533-543.
- Inserm (dir.). *Psychothérapie : Trois approches évaluées*. Rapport. Paris : Les éditions Inserm, 2004, XII- 553 p. - (Expertise collective). - <http://hdl.handle.net/10608/146>.
- Ireps, site web novembre 2012 -accédé le 21/03/2017 : <http://ireps.oxatis.com/Files/93216/142081495446993.pdf>.

- Jacobson, N. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 1.).
- Janse, P., Boezen-Hilberdink, L., van, D. M. K., Verbraak, M. J. P. M., & Hutschemaekers, G. J. M. (2014). Measuring Feedback From Clients: The Psychometric Properties of the Dutch Outcome Rating Scale and Session Rating Scale. *European Journal of Psychological Assessment*, 30, 2, 86-92.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). Big Five Inventory. PsycTESTS Dataset. doi:10.1037/t07550-000
- Jospe, K., Flöel, A., & Lavidor, M. (2018). The interaction between embodiment and empathy in facial expression recognition. *Social Cognitive and Affective Neuroscience*, 13, 2, 203-215.
- Keltner, D., & Haidt, J. (2001). Social functions of emotions. In T. Mayne & G. A. Bonanno (Eds.), *Emotions: Current issues and future directions*. New York: Guilford Press. (pp. 192-213).]
- Kendrick T., El-Gohary M., Stuart B., Gilbody S., Churchill R., Aiken L., Bhattacharya A., Gimson A., Brütt AL., de Jong K., Moore M.. Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *Cochrane Database of Systematic Reviews* 2016, Issue 7. Art. No.: CD011119. DOI: 10.1002/14651858.CD011119.pub2.
- Kessler, R. G., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Zaslavsky, A. M. (2003). Screening for Serious Mental Illness in the General Population. *Archives of General Psychiatry*, 60, 2, 184.
- Kingsley, C., Patel, S., (2017). Patient-reported outcome measures and patient-reported experience measures, *BJA Education*, Volume 17, Issue 4, Pages 137–144, <https://doi.org/10.1093/bjaed/mkw060>.
- Knapp, M. L. (1972). *Nonverbal communication in human interaction*. New York: Holt Rinehart and Winson.

- Knaup, C., Koesters, M., Schoefer, D., Becker, T., & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. *The British Journal of Psychiatry*, 195, 15-22.
- Krägeloh, C. U., Czuba, K. J., Billington, D. R., Kersten, P., & Siegert, R. J. (2015). Using Feedback From Patient-Reported Outcome Measures in Mental Health Services: A Scoping Study and Typology. *Psychiatric Services*, 66, 3, 224-241.
- Kraft, S., Puschner, B., Lambert, M. J., & Kordy, H. (2006). Medical utilization and treatment outcome in mid- and long-term outpatient psychotherapy. *Psychotherapy Research*, 16(2), 241–249.
- Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine*. 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x.
- Kupfer, D. J., First, M. B., & Regier, D. A. (2002). *A research agenda for DSM-V*. Washington, D.C: American Psychiatric Association.
- Krageloh 2015, Krageloh CU, Czuba KJ, Billington DR, Kersten P, Siegert RJ. Using feedback from patient-reported outcome measures in mental health services: a scoping study and typology. *Psychiatric Services* 2015;66(3) :224–41.
- Lambert, M. J. (2010). "Yes, it is time for clinicians to routinely monitor treatment outcome" In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (pp. 239-266). Washington, DC, US: American Psychological Association. <http://dx.doi.org/10.1037/12075-008>.
- Lambert, M. J. (2013). *Bergin and Garfield's handbook of psychotherapy and behavior change*. John Wiley & Sons
- Lambert, M. J. (2015). Progress feedback and the OQ-system: The past and the future. *Psychotherapy (chicago, Ill.)*, 52, 4, 381-90.
- Lambert, M. J. (2017). Maximizing Psychotherapy Outcome beyond Evidence-Based Medicine. *Psychotherapy and Psychosomatics*, 86, 2, 80-89.

- Lambert, M.J., & Burlingame, G.M. (1996a). Outcome Questionnaire (OQ 45.2). Stevenson, NM: American Professional Credentialing Services.
- Lambert, M.J., & Burlingame, G.M. (1996b). Outcome Questionnaire (OQ HAI-22). Stevenson, NM: American Professional Credentialing Services.
- Lambert, M. J., & Hansen, N. B. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69, 2, 159-172.
- Lambert, M. J. & Ogles, B. M. (1997). The effectiveness of psychotherapy supervision. In C. E. Watkins (Ed.), *Handbook of Psychotherapy Supervision* (pp. 421–446). New York: Wiley.
- Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139-193). New York: Wiley.
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L. and Smart, D. W. (2003), Is It Time for Clinicians to Routinely Track Patient Outcome? A Meta-Analysis. *Clinical Psychology: Science and Practice*, 10: 288-301. doi:[10.1093/clipsy.bpg025](https://doi.org/10.1093/clipsy.bpg025).
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy* (chicago, Ill.), 48, 1, 72-9.
- Lambert, M., Whipple, J., Smart, D., Vermeersch, D., Nielsen, S., & Hawkins, E. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy. *Psychotherapy Research*, 11, 49-68.
- Laraway, C,J, 2015. The therapeutic alliance, ruptures, and session-by-session feedback. A thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Clinical Psychology, validated by the University of Oxford (Trinity Term)
- Li X., Hong X., Moilanen A., Huang X., Pfister T., Zhao G., Pietikäinen M. (2015). Reading Hidden Emotions: Spontaneous Micro-expression Spotting and Recognition.

- Lilienfeld, S. O. (2007). Psychological Treatments That Cause Harm. *Perspectives on Psychological Science*, 2, 1, 53-70.
- Marshall M, Lockwood A, Green G, Zajac-Roles G, Roberts C, Harrison G. Systematic assessments of need and care planning in severe mental illness: cluster randomised controlled trial. *Br J Psychiatry* 2004; 185: 163–8.
- Matsumoto, D., Frank, M. & Hwang, H. (2013). Reading people: introduction to the world of nonverbal behavior. In D. Matsumoto M. G. Frank & H. S. Hwang (Eds.), *Nonverbal communication: Science and applications* (pp. 3-14). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781452244037.n1
- Matsumoto, D. & Hwang, H. (2013). Facial expressions. In D. Matsumoto M. G. Frank & H. S. Hwang (Eds.), *Nonverbal communication: Science and applications* (pp. 15-52). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781452244037.n2
- Matsumoto, D. & Hwang, H. (2013). Cultural influences on nonverbal behavior. In D. Matsumoto M. G. Frank & H. S. Hwang (Eds.), *Nonverbal communication: Science and applications* (pp. 97-120). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781452244037.n5
- Matsumoto, D. Willingham, B. (2006). The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners of the 2004 athens olympic games. *Journal of Personality and Social Psychology*, 91, 3, 2006.
- Minami, T., Brown, G. S., McCulloch, J., & Bolstrom, B. J. (2012). Benchmarking therapists: furthering the benchmarking method in its application to clinical practice. *Quality and Quantity*, 46, 6, 1699-1708.
- Miller, Scott. (2006). Miller, S.D., Duncan, B.L., Sorrell, R., Brown, G.S., & Chalk, M.B. (2006). *Using outcome to inform therapy practice. Journal of Brief Therapy*, 5(1), 5-22..
- Miller, S. D., & Duncan, B. L. (2000, 2004). *The Outcome and Session Rating Scales*. ISTC, Chicago

- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, 2(2), 91–100.
- Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology*, 61, 2, 199-208.
- Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The outcome of psychotherapy: Yesterday, today, and tomorrow. *Psychotherapy*, 50, 1, 88-97.
- Miller, S.D., Hubble, M.A., & Duncan, B.L. (2007). Supershrinks: Learning from the field's most effective practitioners. *The Psychotherapy Networker*, 31(6), 26-35, 56..
- Miller, Scott & A Hubble, Mark & Chow, Daryl & Seidel, Jason. (2015). Beyond Measures and Monitoring: Realizing the Potential of Feedback-Informed Treatment. *Psychotherapy (Chicago, Ill.)*. 52. 449-457. 10.1037/pst0000031.
- Miller, S. D., Wampold, B. E., & Varhely, K. (2008). Direct comparisons of treatment modalities for youth disorders: A meta-analysis. *Psychotherapy Research*, 18, 5-14.
- Moggia, D., Niño-Robles, N., Miller, S.D., Feixas, G. (2017). Psychometric Properties of the Outcome Rating Scale (ORS) and Session Rating Scale 3.0 (SRS 3.0) in a Spanish Clinical Sample. In F. Giannone (Chair). Brief Paper Session: Quantitative & Qualitative Method. In Society for Psychotherapy Research UK & European Chapters 4th joint conference, Oxford.
- Moggia, D., Niño-Robles, N., Miller, S. D., & Feixas, G. (2018). Psychometric Properties of the Outcome Rating Scale (ORS) in a Spanish Clinical Sample. *The Spanish Journal of Psychology*, 21.
- Muran, J. C., Safran, J. D., Gorman, B. S., Samstag, L. W., Eubanks-Carter, C., & Winston, A. (2009). The relationship of early alliance ruptures and their resolution to process and outcome in three time-limited psychotherapies for personality disorders. *Psychotherapy: Theory, Research, Practice, Training*, 46, 2, 233-248.
- Muthén L.K, & Muthén B.O (2010). Mplus user's guide. Los Angeles, CA: Muthén & Muthén.

- Nguyen, L., Blasquez, S., Bataille, B., & Chassery, C. (2012). La détresse psychologique mesurée par le score de Kessler (K6) prédit les douleurs postopératoires prolongées après chirurgie du poignet. *Canadian Journal of Anesthesia/Journal Canadien D'anesthésie*, 59(12), 1150–1151. doi:10.1007/s12630-012-9783-8
- Niedenthal, Paula M, and Francois Ric. *Psychology of Emotion*. New York [etc.: Routledge, 2017. Print.
- Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2003). Fifty years of process-outcome research: Continuity and change. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 307-390). New York: Wiley.
- Orlinsky, D.E., Romnestad, M.H, Willutzki, U. CH8 in Lambert, M. J. (2004). *Bergin and Garfield's handbook of psychotherapy and behavior change*. New York, NY: Wiley.
- Østergård, O. K., Randa, H., & Hougaard, E. (2018). The effect of using the Partners for Change Outcome Management System as feedback tool in psychotherapy-A systematic review and meta-analysis. *Psychotherapy Research : Journal of the Society for Psychotherapy Research*, 1-18.
- Plaisant, O., Courtois, R., Réveillère, C., Mendelsohn, G. A., & John, O. P. (2010). Validation par analyse factorielle du Big Five Inventory français (BFI-Fr). Analyse convergente avec le NEO-PI-R. *Annales Médico-Psychologiques, Revue Psychiatrique*, 168(2), 97–106. doi:10.1016/j.amp.2009.09.003
- Pittam, J. (1994). *Voice in social interaction: An interdisciplinary approach*. Thousand Oaks, Calif: Sage Publications.
- Porter, S., & ten, B. L. (2008). Reading Between the Lies: Identifying Concealed and Falsified Emotions in Universal Facial Expressions. *Psychological Science*, 19, 5, 508-514.
- Pringle, J., & Coley, K. (2015). Improving medication adherence: a framework for community pharmacy-based interventions. *Integrated Pharmacy Research and Practice*, 175. doi:10.2147/iprp.s93036.

- Pringle J, Melczak M, Aldridge A, et al. (2011). Medication adherence and its relationship to the therapeutic alliance: Results from an innovative pilot study within a community pharmacy MTM practice. *Inov Pharm.*;2(1): Article 33. <http://pubs.lib.umn.edu/innovations/vol2/iss1/10>.
- Quirk, K., Miller, S., Duncan, B., & Owen, J. (2013). Group Session Rating Scale: Preliminary psychometrics in substance abuse group interventions. *Counselling and Psychotherapy Research*, 13 (3), 194–200. doi:10.1080/14733145.2012.744425
- Reese, R. J., Gillaspay, J. A., Owen, J. J., Flora, K. L., Cunningham, L. C., Archie, D., & Marsden, T. (2013). The Influence of Demand Characteristics and Social Desirability on Clients' Ratings of the Therapeutic Alliance. *Journal of Clinical Psychology*, 69 (7), 696–709. doi:10.1002/jclp.21946
- Reese, R. J., Toland, M. D., Slone, N. C., & Norsworthy, L. A. (2010). Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy (chicago, Ill.)*, 47, 4, 616-30.
- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology*, 51(3), 649-660. <http://dx.doi.org/10.1037/0022-3514.51.3.649>
- Roter, D. L., Frankel, R. M., Hall, J. A., & Sluyter, D. (2006). The expression of emotion through nonverbal behavior in medical visits: Mechanisms and outcomes. *Journal of General Internal Medicine*, 21, 1, 28-34.
- Safran, J. & Muran, J., & Eubanks, C. (2011). Repairing Alliance Ruptures. *Psychotherapy relationships that work: Evidence-based responsiveness*. 48. 224-238. 10.1093/acprof:oso/9780199737208.003.0011.
- Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology*, 61, 2, 145-153.
- Schlegel, K., Grandjean, D., & Scherer, K. R. (2014). Introducing the Geneva Emotion Recognition Test: An example of Rasch-based test development. *Psychological Assessment*, 26, 2, 666-672.

- Schlegel, K., & Scherer, K. R. (2016). Introducing a short version of the Geneva Emotion Recognition Test (GERT-S): Psychometric properties and construct validation. *Behavior Research Methods*, 48, 4, 1383-1392.
- Schmidt U, Landau S, Pombo-Carril MG, Bara-Carril N, Reid Y, Murray K, et al. Does personalized feedback improve the outcome of cognitive-behavioural guided self-care in bulimia nervosa? A preliminary randomized controlled trial. *Brit J Clin Psychol* 2006; 45: 111–21.
- Schuckard, E., & Miller, S. D. (2016). *Psychometrics of the ORS and SRS. Results from RCTs and Meta-analyses of Routine Outcome Monitoring & Feedback. The Available Evidence.* Chicago, IL. <https://www.scottdmiller.com/wp-content/uploads/2016/09/Measures-and-Feedback-2016.pdf>.
- Seidel, J. A., Andrews, W. P., Owen, J., Miller, S. D., & Buccino, D. L. (2016). Preliminary validation of the Rating of Outcome Scale and equivalence of ultra-brief measures of well-being. *Psychological Assessment*, 29 (1), 65–75. doi:10.1037/pas0000311
- Shadish, W. R., Matt, G. E., Navarro, A. M., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, 126, 512–529.
- Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy*. Vol. 15, No. 5, Pg. 287-303
- Slade M, McCrone P, Kuipers E, Leese M, Cahill S, Parabiaghi A, et al. Use of standardised outcome measures in adult mental health services. Randomised controlled trial. *Br J Psychiatry* 2006; 189: 330–6.
- Smith, S., Crocker, L., Staton, C., Gillaspay, A., Charlton, S Ph.D. (2010) Psychometric Properties of the Outcome Rating Scale in a Non-Clinical Sample, insepified date.. Poster. University of Central Arkansas. Retrieved on the internet the 19/04/2018:
- Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Englewood Cliffs, NJ: Prentice-Hall.

- Spitzer, R. L., Kroenke, K., Williams, J. B., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166, 1092-1097. doi:10.1001/archinte.166.10.1092
- Stiles, W. B., Barkham, M., Connell, J., & Mellor-Clark, J. (2008). Responsive Regulation of Treatment Duration in Routine Practice in United Kingdom Primary Care Settings: Replication in a Larger Sample. *Journal of Consulting and Clinical Psychology*, 76, 2, 298-305.
- Stiles, W. B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008). Effectiveness of cognitive-behavioural, person-centred, and psychodynamic therapies in UK primary care routine practice: *Replication in a larger sample*. *Psychological Medicine*, 38, 677–68.
- Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of Early Change in Psychotherapy under Routine Outpatient Conditions. *Journal of Consulting and Clinical Psychology*, 75, 6, 864-874.
- Sugarman, M. A. (2016). Are antidepressants and psychotherapy equally effective in treating depression? A critical commentary. *Journal of Mental Health (abingdon, England)*, 25, 6, 475-478.
- Sutter, JV., "Assessing impact of affect recognition on the peutic relationship" (2010). University of Kentucky Master's theses. [ps://uknowledge.uky.edu/gradschool_theses/14](https://uknowledge.uky.edu/gradschool_theses/14)
- Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80, 547–559.
- Tracey TJ, Kokotovic AM. Factor structure of the working alliance inventory. *Psych Assess* 1989; 1 (3): 207–10.
- Trudeau LS. Effects of a Clinical Feedback System on Client and Therapist Outcomes in a Rural Community Mental Health Center. Doctoral Thesis 21 (PhD), Institute for Social and Behavioral Research, Iowa State University, USA, 2001.

- Umberson, D., & Montez, J. K. (2010). Social Relationships and Health: A Flashpoint for Health Policy. *Journal of Health and Social Behavior*, 51(Suppl), S54–S66. <http://doi.org/10.1177/0022146510383501>
- Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la recherche en langue française [Toward a methodology for the transcultural validation of psychological questionnaires: Implications for research in the French language]. *Canadian Psychology/Psychologie canadienne*, 30(4), 662-680. <http://dx.doi.org/10.1037/h0079856>
- Wampold, B. E. (2001a). Contextualizing Psychotherapy as a Healing Practice: Culture, History, and Methods. *Applied and Preventive Psychology*, 10, 69-86.
- Wampold, B. E. (2001b). *The great psychotherapy debate: Model, methods, and findings*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York. Routledge.
- Wampold, B. E., Minami, T., Baskin, T. W., & Tierney, S. C. (2002). A meta-re(analysis) of the effects of cognitive therapy versus “other therapies” for depression. *Journal of Affective Disorders*, 68, 159-165.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, “all must have prizes.” *Psychological Bulletin*, 122 (3), 203-215.
- Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G. A. (2010). Youth psychotherapy change trajectories and outcomes in usual care: Community mental health versus managed care settings. *Journal of Consulting and Clinical Psychology*, 78(2), 144-155
- Warren, G., Schertler, E., & Bull, P. (2009). Detecting Deception from Emotional and Unemotional Cues. *Journal of Nonverbal Behavior*, 33, 1, 59-69.

- Wellman, H. M. (2018). Theory of mind across the lifespan? *Zeitschrift für Psychologie*, 226(2), 136-138. <http://dx.doi.org/10.1027/2151-2604/a000330>.
- Werbart, A., Levin, L., Andersson, H., & Sandell, R. (2013). Everyday evidence: Outcomes of psychotherapies in Swedish public health services. *Psychotherapy*, 50, 1, 119-130.
- Whipple, J., Lambert, M., Vermeersch, D., Smart, D., Nielsen, S. & Hawkins, E. (2003). Improving the Effects of Psychotherapy. *Journal of Counseling Psychology*, 50(1), 59-68.
- Wierzbicki, M., Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice*, Vol 24(2), 190-195.
- Worthen, V. E., & Lambert, M. J. (2007). Outcome oriented supervision: Advantages of adding systematic client tracking to supportive consultations. *Counselling and Psychotherapy Research*, 7, 1, 48-53.

5. ANNEXES

ANNEXE 1 : POSTER PRESENTE AU CONGRES DE L'AFTCC EN 2014	232
ANNEXE 2 : ECHELLE D'EVALUATION DES RESULTATS DE LA SEANCE (ORS ; MILLER & AL., 2005)	233
ANNEXE 3 : ECHELLE D'EVALUATION DES RESULTATS DE LA SEANCE (SRS ; MILLER & AL., 2005)	234
ANNEXE 4 : SEQUENCE DE RECUEIL DES DONNEES (ETUDE VFERA)	235
ANNEXE 5 : CONSENTEMENT ECLAIRE DES THERAPEUTES (VFERA)	238
ANNEXE 6 : CONSENTEMENT ECLAIRE DES PATIENTS (VFERA)	239
ANNEXE 7 : CONSENTEMENT ECLAIRE DES ETUDIANTS (VFERA)	240
ANNEXE 8 : BROCHURE D'INSCTRUCTIONS POUR LES THERAPEUTES (VFERA)	241
ANNEXE 9 : FICHE DE POINTAGE (VFERA)	253
ANNEXE 10 : BROCHURE D'INSTRUCTION POUR LES ETUDIANTES (VFERA)	254
ANNEXE 11 : BROCHURE D'INSTRUCTIONS POUR LES THERAPEUTES ANGLOSAXONS (NVB-O)	262
ANNEXE 12 : FORMULAIRE DE CONSENTEMENT ECLAIRE (NVB-O)	265

ANNEXE 1 : POSTER PRESENTE AU CONGRES DE L'AFTCC EN 2014

Christophe Cazauvielh – Psychologue
christophe.cazauvielh@u-bordeaux.fr
Laboratoire de Psychologie Santé et qualité de vie – EA 4139 Université Victor Segalen Bordeaux 2 - 3 ter Place de la Victoire - 33076 Bordeaux Cedex

université
BORDEAUX

Evaluer et améliorer l'efficacité des psychothérapies : revue de la littérature et présentation d'un système de suivi des résultats en continu (ROM)

Problématique et enjeux

Au premier plan des sujets de santé publique, figurent l'efficacité des psychothérapies et leur encadrement. Bien que l'efficacité générale des psychothérapies soit réputée importante (Lambert et Ogles, 2004) :

- Une proportion non négligeable de patients manifestent une détérioration pendant leur traitement (5 à 10% ; Cooper, 2008)
- Les taux d'abandon en thérapie sont élevés (jusqu'à 47% ; Wierzbicki et Pekarik, 1993)
- La variation d'efficacité entre les thérapeutes est importante. Les thérapeutes ont des difficultés à détecter les patients qui sont à risque de se détériorer ou d'abandonner la thérapie (Lambert et Ogles, 1997)
- La possibilité d'obtenir des données probantes en pratique clinique est souvent freinée par la lourdeur des protocoles. Ces données seraient susceptibles de guider la délivrance des soins, de documenter l'efficacité des traitements en vue d'un remboursement, et d'aider les cliniciens dans leur développement professionnel continu.

10% des patients manifestent une détérioration pendant leur traitement
47% de taux d'abandon en psychothérapie

Revue de la littérature – Intérêt du suivi continu des résultats

L'alliance thérapeutique et les résultats précoces sont associés aux meilleurs résultats

- 2 indicateurs sont associés aux meilleurs résultats psychothérapeutiques :
 - Une bonne alliance thérapeutique entre le clinicien et son patient (Baldwin et al., 2007)
 - Des résultats positifs sur les premières séances de traitement (Miller et al., 2006)

Le système trans-théorique et trans-diagnostic de suivi en continu des résultats *Partners for Change Outcome Management System* (PCOMS), permet de recueillir le feedback des patients sur ces 2 indicateurs, et d'utiliser ces informations pour guider la délivrance des soins :

- 1) L'échelle Outcome Rating Scale (ORS, Miller et al., 2003) mesure le changement thérapeutique
- 2) L'échelle Session Rating Scale (SRS, Duncan et al., 2003) mesure l'alliance thérapeutique

Ces échelles ont été reconnues valides, facilement utilisables en pratique clinique, et sont inscrites au répertoire américain des programmes et pratiques empiriquement validés (NREPP)

Une compilation de la littérature sur l'efficacité des différents systèmes de suivi des résultats et de feedback (Miller et Schuckard , 2014) documente :

- Une diminution des taux d'abandon (jusqu'à 50%)
- Une diminution des taux de détérioration pendant le traitement (jusqu'à 33%)
- Une augmentation du taux de changement cliniquement significatif (jusqu'à 2 fois), ainsi qu'une baisse des coûts de santé

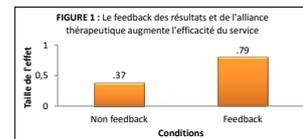
Diminution de **50%** des taux d'abandon
Baisse de **33%** de la détérioration

Quelques résultats spécifiques à l'utilisation du feedback avec PCOMS

Etude 1 : Service d'assistance téléphonique aux employés (Etats-Unis)

Etude quasi-expérimentale (Miller et al., 2006 ; n=6424, 66% femmes, âge moyen=36, $\sigma=13$), clients rencontrant des problématiques typiques d'une clinique de soin de santé mentale (anxiété, dépression, abus de substances...).

➔ Amélioration significative des résultats des services entre la ligne de base (condition non feedback) et l'introduction du feedback (figure 1)



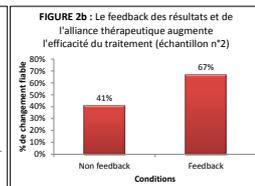
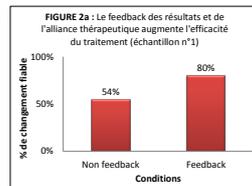
Etude 2 : Service de counseling universitaire et centre de thérapie familiale universitaire (Etats-Unis)

Etude contrôlée randomisée (Reese et al., 2009) comparant l'efficacité d'un traitement habituel (TAU) et la condition de délivrance du feedback.

Echantillon n°1 : clients recevant une thérapie dans un centre de counseling universitaire (n=74, 53 femmes, 18 hommes, 3 indéterminés, âge 18-27, $m=17$, $\sigma=1.90$).

Echantillon n°2 : clients recevant les soins d'une clinique de formation de thérapies familiales (n=74, 51 femmes, 21 hommes, 2 indéterminés, âge 18-69, âge moyen=32,96, $\sigma=12,32$).

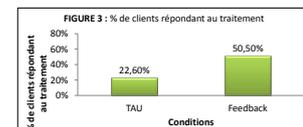
➔ Augmentation significative des gains dus au traitement pour les clients dans la condition feedback (figures 2a et 2b)



Etude 3 : Thérapies de couple (Norvège)

Etude contrôlée randomisée (Anker et al., 2009 ; 410 couples hétérosexuels, âge 20-71, $m=37,83$, $\sigma=8,48$), comparant l'efficacité d'un traitement habituel (TAU) et la condition de délivrance du feedback, dans une agence de conseil familial financée par le gouvernement.

➔ Augmentation sensible du changement fiable ou cliniquement significatif dans la condition feedback (Figure 3)



Conclusion et perspectives de recherche

L'utilisation d'un système de gestion des résultats améliore la qualité et l'efficacité des soins délivrés. L'adaptation française du système PCOMS en collaboration avec le Dr Miller est en cours (Christophe Cazauvielh), et permettra d'amener des éléments concrets de réponse aux enjeux d'évaluation et d'encadrement des pratiques psychothérapeutiques.

Références bibliographiques et participation à la recherche (psychiatres, psychologues)
Contact : christophe.cazauvielh@u-bordeaux.fr ou christophe.cazauvielh@gmail.com

ANNEXE 2 : ECHELLE D'ÉVALUATION DES RESULTATS DE LA SEANCE (ORS ; MILLER & AL., 2005)

ÉCHELLE D'ÉVALUATION DES RÉSULTATS (ORS)

Nom _____ Âge (années) : _____ Sexe : _____
Séance N° _____ Date : _____
Pour qui remplissez-vous ce formulaire ? Mettre une croix s'il vous plaît : Pour vous-même <input type="checkbox"/> Pour autrui <input type="checkbox"/>
Si Pour autrui, quel lien avez-vous avec cette personne ? _____

Considérez la semaine qui vient de s'écouler (ou la période écoulée depuis votre dernière visite), y compris aujourd'hui ; aidez-nous à comprendre comment vous vous êtes senti(e) en évaluant votre bien-être dans les aspects suivants de votre vie. L'extrémité gauche de l'échelle représente le plus bas niveau de bien-être et l'extrémité droite de l'échelle représente le plus haut niveau de bien-être. *Si vous remplissez ce formulaire pour quelqu'un d'autre, faites-le s'il vous plaît en fonction de ce que vous pensez de son bien-être.*

À L'ATTENTION DU CLINICIEN : AFIN D'ASSURER L'EXACTITUDE DE LA COTATION, IMPRIMEZ CETTE PAGE ET VÉRIFIEZ QUE LES ECHELLES (LES LIGNES DE POINTILLÉS) MESURENT BIEN 10 CM DE LONG. SI CE N'EST PAS LE CAS, MODIFIEZ LE FORMULAIRE JUSQU'À CE QUE LES LIGNES S'IMPRIMENT À LA LONGUEUR CORRECTE. ENSUITE, EFFACEZ CE MESSAGE.

Individuellement

(Bien-être personnel)

I-----I

Interpersonnel

(Famille, relations proches)

I-----I

Socialement

(Travail, école, amitiés)

I-----I

Globalement

(Sentiment général de bien-être)

I-----I

International Center for Clinical Excellence

www.scottdmiller.com

© 2000, Scott D. Miller and Barry L. Duncan

Traduction : François Simard et John Deltour. Révision : Christophe Cazauvielh
(christophe.cazauvielh@gmail.com), John Deltour ([johndeltour@yahoo.fr](mailto: johndeltour@yahoo.fr)), Raymond Lemay ([rlemay@valorispr.ca](mailto: rlemay@valorispr.ca)), Chantal Tassé
([ctasse@valorispr.ca](mailto: ctasse@valorispr.ca)), Madeleine Lalonde ([malalonde@valorispr.ca](mailto: malalonde@valorispr.ca)), Nicholas Cardinal ([nicardinal@valorispr.ca](mailto: nicardinal@valorispr.ca))

Licence d'utilisation individuelle uniquement

ANNEXE 3 : ECHELLE D'ÉVALUATION DES RESULTATS DE LA SEANCE (SRS ; MILLER & AL., 2005)

ÉCHELLE D'ÉVALUATION DE LA SÉANCE (SRS V.3.0)

Nom _____ Âge (années): _____
Sexe : _____
Séance N° _____ Date : _____

S'il vous plaît, évaluez la séance d'aujourd'hui en plaçant un trait sur la ligne le plus près de la description correspondant le mieux à votre expérience.

	Relation	
Je ne me suis pas senti(e) écouté(e), compris(e), et respecté(e).	-----	Je me suis senti(e) écouté(e), compris(e), et respecté(e).
	Objectifs et Thèmes	
Nous n'avons pas travaillé sur ce que je voulais ou n'avons pas parlé de ce dont je voulais.	-----	Nous avons travaillé sur ce que je voulais et nous avons parlé de ce dont je voulais.
	Approche ou Méthode	
L'approche du thérapeute ne me convient pas.	-----	L'approche du thérapeute me convient.
	Globalement	
Il manquait quelque chose dans la séance d'aujourd'hui.	-----	Globalement, la séance d'aujourd'hui m'a convenu.

International Center for Clinical Excellence

www.scottdmiller.com

© 2002, Scott D. Miller, Barry L. Duncan, & Lynn Johnson

Traduction: Serge Galarnau et John Deltour. Révision : Christophe Cazauvielh (christophe.cazauvielh@gmail.com), John Deltour (johndeltour@yahoo.fr), Raymond Lemay (rlemay@valorispr.ca), Chantal Tassé (ctasse@valorispr.ca), Madeleine Lalonde (malalonde@valorispr.ca), Nicholas Cardinal (nicalcardinal@valorispr.ca)

Licence d'utilisation individuelle uniquement

Validation des échelles ORS/SRS

Données enregistrées

Echantillon Clinique 1

Condition Non-Feedback- administration en ligne

2 échantillons étudiants

Condition Non-Feedback - administration en ligne

Séance 1 Consentement éclairé	Séance 1 Consentement éclairé -test-
Données socio-démographiques (Age, Sexe, Statut Matrimonial , Niveau étude, CSP)	Données socio-démographiques (Age, Sexe, Statut Matrimonial , Niveau étude, CSP)
	Psychothérapie & Traitement médicamenteux (O/N)
ORS (Début de la séance)	ORS
Kessler K-6, MI-22	Kessler K-6
	MI-22
SRS, WAI-S (Fin de la séance)	BPI

Séance 2 (Délais variables)	Re-Test (+14 jours)
ORS/SRS	ORS/MI-22/K6 Stop enregistrement

Séance 3 (Délais variables)	
ORS/SRS MI-22	

Séances 4 et 5	
ORS/SRS	

Séance 6	
ORS/MI-22 SRS	
Séances suivantes (sauf S10 + MI-22) ORS/SRS	

2ème site - Echantillon clinique (pratique privée)

Condition Feedback- administration en ligne (patients qui ont commencé leur traitement avec la version unifiée de l' ORS/SRS).

Sexe

Age,

Statut Matrimonial,

Profession (O/N)

Evaluation psychologique (Problème principal)

Traitement Médicamenteux connu ()

Les questionnaires ORS et SRS sont administrés en continu

ANNEXE 5 : CONSENTEMENT ECLAIRE DES THERAPEUTES (VFERA)

Consentement éclairé

Cette recherche doctorale sur l'évaluation des Psychothérapies, nommée Etude VFERA, est menée par Christophe Cazauvielh et s'inscrit dans le cadre des travaux du Laboratoire de Psychologie, Santé et Qualité de vie EA 4139 de l'université de Bordeaux.

Cette recherche est conforme aux dispositions légales en vigueur et respecte les principes fixés par le code de déontologie des Psychologues, ainsi que ceux fixés par le Laboratoire de Psychologie, Santé et Qualité de vie EA 4139.

- Le responsable de la recherche vous garantit la confidentialité des données recueillies.
- En tant que thérapeute vous n'aurez pas accès aux données recueillies auprès des patients et vous vous engagez à tenir confidentiels les codes d'accès aux questionnaires associés à chaque patient. Vous vous engagez par ailleurs à ne pas utiliser les questionnaires ORS/SRS/MI-22 avec vos patients qui participent à cette recherche, puisqu'ils sont inclus dans le protocole.
- Les données ne seront utilisées que dans le cadre de la recherche.
- La qualité des résultats de cette recherche dépend de votre participation, toutefois vous avez le droit de vous retirer de cette recherche quand vous le souhaitez, et vous pouvez à tout moment, demander à accéder à vos données, et exercer un droit de rectification auprès du responsable de la recherche aux coordonnées suivantes :

Christophe CAZAUVIEILH
Laboratoire de Psychologie, Santé et Qualité de vie EA4139
3 ter place de la Victoire
33076 Bordeaux Cedex

- Vous pourrez connaître les résultats généraux de l'étude lorsqu'ils seront disponibles, si vous le désirez.
- Votre participation repose sur le volontariat et nécessite la signature datée de ce consentement éclairé, en double exemplaire, dont un exemplaire à conserver sans délais et un exemplaire à renvoyer au responsable de cette recherche.

J'ai pris connaissance des implications de cette recherche et je signe mon consentement éclairé.

A Le...../...../.....

Nom..... Prénom.....

Profession.....

SIGNATURE

Exemplaire à conserver par le thérapeute

1/2

ANNEXE 6 : CONSENTEMENT ECLAIRE DES PATIENTS (VFERA)



DEBUT DE LA SEANCE - VFERA

Informations sur la recherche et consentement éclairé

Cette recherche doctorale nommée Etude VFERA, est menée par Christophe Cazauviel, Psychologue, et s'inscrit dans le cadre des travaux du Laboratoire de Psychologie, Santé et Qualité de vie EA 4139 de l'université de Bordeaux. Les données issues de cette recherche permettront de mieux évaluer l'efficacité des Psychothérapies.

A cet effet, votre thérapeute vous demandera de remplir des questionnaires visant à mesurer votre bien-être ainsi qu'à évaluer votre perception de la thérapie, lors de votre première consultation, puis lors des séances suivantes.

Cette recherche est conforme aux dispositions légales en vigueur et respecte les principes fixés par le Code de déontologie des Psychologues, ainsi que ceux fixés par le Laboratoire de Psychologie, Santé et Qualité de vie EA 4139.

- Le responsable de la recherche vous garantit la confidentialité des données recueillies.
- Votre thérapeute n'aura pas accès aux données des questionnaires.
- Les données ne seront utilisées que dans le cadre de la recherche.

• La qualité des résultats de cette recherche dépend de votre participation, toutefois vous avez le droit de vous retirer de cette recherche quand vous le souhaitez, et vous pouvez à tout moment demander à accéder à ces données et exercer un droit de rectification auprès du responsable de la recherche aux coordonnées suivantes :

Christophe CAZAUVIELH
Laboratoire de Psychologie, Santé et Qualité de vie EA4139
– 3 ter place de la Victoire – 33076 Bordeaux Cedex
christophe.cazauvielh@u-bordeaux.fr

- Vous pourrez connaître les résultats généraux de l'étude lorsqu'ils seront disponibles, si vous le désirez. Nous vous remercions par avance pour votre participation à cette recherche.
- Votre participation repose sur le volontariat et nécessite la signature électronique datée de ce consentement éclairé.

ANNEXE 7 : CONSENTEMENT ECLAIRE DES ETUDIANTS (VFERA)



Informations sur la recherche et consentement éclairé

Cette recherche doctorale nommée Etude VFERA, est menée par Christophe Cazauviel, Psychologue, et s'inscrit dans le cadre des travaux du Laboratoire de Psychologie, Santé et Qualité de vie EA 4139 de l'université de Bordeaux. Les données issues de cette recherche permettront de mieux évaluer l'efficacité des Psychothérapies.

A cet effet, des questionnaires évaluant votre bien-être et votre personnalité vont vous être présentés. Cette recherche s'effectuera en deux temps : une première série de questions vous sera présentée aujourd'hui et une seconde série à 15 jours, de date à date.

Cette recherche est conforme aux dispositions légales en vigueur et respecte les principes fixés par le Code de déontologie des Psychologues, ainsi que ceux fixés par le Laboratoire de Psychologie, Santé et Qualité de vie EA 4139.

- Le responsable de la recherche vous garantit la confidentialité des données recueillies et du traitement de ces données.
- Les données ne seront utilisées que dans le cadre de la recherche. Si par ailleurs vous pensez que vous avez besoin d'un soutien psychologique, vous pouvez contacter directement votre médecin traitant.
- La qualité des résultats de cette recherche dépend de votre participation, toutefois vous avez le droit de vous retirer de cette recherche quand vous le souhaitez, et vous pouvez à tout moment demander à accéder à ces données et exercer un droit de rectification auprès du responsable de la recherche, aux coordonnées suivantes :

Christophe CAZAUVIELH
Laboratoire de Psychologie, Santé et Qualité de vie EA4139
– 3 ter place de la Victoire – 33076 Bordeaux Cedex
christophe.cazauviel@u-bordeaux.fr

- Vous pourrez connaître les résultats généraux de l'étude lorsqu'ils seront disponibles, si vous le désirez. Nous vous remercions par avance pour votre participation à cette recherche.
- Votre participation repose sur le volontariat et nécessite la signature électronique datée de ce consentement éclairé.



En cochant la case suivante, je spécifie que j'accepte de participer à l'étude VFERA

J'accepte de participer à l'étude VFERA et je signe mon consentement éclairé.

université
de BORDEAUX

ETUDE VFERA

INSTRUCTIONS POUR LE THERAPEUTE

Contact : Christophe CAZAUVIEILH

Laboratoire de Psychologie, Santé et Qualité de vie EA4139

– 3 ter place de la Victoire –

33076 Bordeaux Cedex

Sommaire

I. CRITERES D'INCLUSION ET D'EXCLUSION	3
II. PRE-REQUIS TECHNIQUES POUR PARTICIPER A L'ETUDE	4
III. FICHE DE POINTAGE : UTILISATION	5
IV. ADMINISTRATION DES QUESTIONNAIRES	6
A. Au tout début de chaque séance :	6
B. A la fin de chaque séance :	7
C. Quand arrêter d'enregistrer les réponses des patients ?	8
V. EN CAS DE PROBLEME, LES INSTRUCTIONS A DONNER AU PATIENT	10

I. Critères d'inclusion et d'exclusion



Critères d'inclusion



- Patient(e)
 - se présentant à une première séance de psychothérapie (psychologue ou psychiatre) ;
 - **ET** âgé(e) de 18 à 70 ans

Critères d'exclusion



- Etats incompatibles avec l'administration de questionnaires (états d'agitation, états psychotiques aigus ...) ;
- Difficultés avec la lecture et la compréhension de la langue française
- Handicaps incompatibles avec l'administration autonome des questionnaires en ligne :
 - Handicap visuel
 - Handicap cognitif et neurologique
 - Handicap moteur (pour l'utilisation de la souris, trackpad, ou tablette)

II. Pré-requis techniques pour participer à l'étude

- Disposer d'une connexion internet.



- Utiliser un navigateur internet compatible.



L'application est compatible avec ces navigateurs et leurs versions ultérieures :

- SAFARI version 8.0.3
- FIREFOX version 33.1.1
- CHROME version 40.0.2214.115
- INTERNET EXPLORER version 11.0.9600.17691

! AVANT LA PREMIERE UTILISATION !

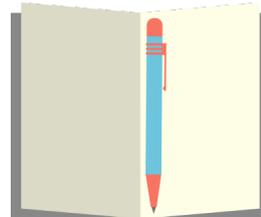
- ⇒ **Mettez à jour votre navigateur internet**
- ⇒ **Redémarrez votre ordinateur**

III. Fiche de Pointage : utilisation

- La fiche de pointage est pré-remplie avec les codes (zone **1**) qui vous permettent de connecter les patients au questionnaire (3 patients par page).

universit� de BORDEAUX		ETUDE VFERA – FICHE DE POINTAGE						
		MERCİ D'ECRIRE EN LETTRES CAPITALES						
INFORMATIONS SUR LE THERAPEUTE								
Age	__ ans	Sexe	1. F <input type="checkbox"/>	2. M <input type="checkbox"/>	Profession	1. Psychologue <input type="checkbox"/>	2. Psychiatre <input type="checkbox"/>	Ann�e(s) d'exp�rience : __ an(s)
PATIENTS								
CODE	Patient mandat�	Evaluation			Approche th�rapeutique utilis�e	Medication psychotrope	Arr�t enregistrement	
1	1. Oui <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/>			-----	1. Oui <input type="checkbox"/>	1. Nombre de s�ances n > 20 <input type="checkbox"/>	
	2. Non <input type="checkbox"/>	2. Trouble anxieux <input type="checkbox"/>					2. Fin de la th�rapie n < 20 <input type="checkbox"/>	
		3. D�pendance et abus de substance <input type="checkbox"/>					3. Abandon de la th�rapie <input type="checkbox"/>	
		4. Trouble psychotique <input type="checkbox"/>					4. D�lais > 8 semaines depuis la s�ance pr�c�dente <input type="checkbox"/>	
		5. D�mence et trouble cognitif s�v�re <input type="checkbox"/>						
		6. Autre : ----- <input type="checkbox"/>						
CODE	Patient mandat�	Evaluation			Approche th�rapeutique utilis�e	Medication psychotrope	Arr�t enregistrement	
	1. Oui <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/>			-----	1. Oui <input type="checkbox"/>	1. Nombre de s�ances n > 20 <input type="checkbox"/>	
	2. Non <input type="checkbox"/>	2. Trouble anxieux <input type="checkbox"/>					2. Fin de la th�rapie n < 20 <input type="checkbox"/>	
		3. D�pendance et abus de substance <input type="checkbox"/>					3. Abandon de la th�rapie <input type="checkbox"/>	
		4. Trouble psychotique <input type="checkbox"/>					4. D�lais > 8 semaines depuis la s�ance pr�c�dente <input type="checkbox"/>	
		5. D�mence et trouble cognitif s�v�re <input type="checkbox"/>						
		6. Autre : ----- <input type="checkbox"/>						
CODE	Patient mandat�	Evaluation			Approche th�rapeutique utilis�e	Medication psychotrope	Arr�t enregistrement	
	1. Oui <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/>			-----	1. Oui <input type="checkbox"/>	1. Nombre de s�ances n > 20 <input type="checkbox"/>	
	2. Non <input type="checkbox"/>	2. Trouble anxieux <input type="checkbox"/>					2. Fin de la th�rapie n < 20 <input type="checkbox"/>	
		3. D�pendance et abus de substance <input type="checkbox"/>					3. Abandon de la th�rapie <input type="checkbox"/>	
		4. Trouble psychotique <input type="checkbox"/>					4. D�lais > 8 semaines depuis la s�ance pr�c�dente <input type="checkbox"/>	
		5. D�mence et trouble cognitif s�v�re <input type="checkbox"/>						
		6. Autre : ----- <input type="checkbox"/>						
Contact : Christophe CAZAUVIELH #cristophe.cazauvieilh@u-bordeaux.fr Laboratoire de Psychologie, Sant� et Qualit� de vie EA4139 – 3 ter place de la Victoire – 33076 Bordeaux Cedex								

- ⇒ **Notez** le code dans votre dossier patient. Cela vous permettra d'acc der au questionnaire lors des s ances suivantes.



IV. Administration des questionnaires

A. Au tout début de chaque séance :

⇒ **Connectez-vous** à internet.



⇒ **Ouvrez** le mail de participation à l'étude



⇒ **Cliquez** sur le lien **DEBUT DE SEANCE –VFERA** (pour mémoire l'adresse est :

⇒ **Vérifiez** que s'affiche bien l'intitulé **DEBUT DE LA SEANCE - VFERA**, dans la fenêtre du navigateur. (Comme affiché ci-dessous)



⇒ **Entrez** le code correspondant à votre patient

⇒ **Cliquez** sur *continuer*

⇒ **Placez le patient devant l'écran pour qu'il remplisse seul le questionnaire.**

⇒ **Indiquez** au patient qu'à la fin de cette série de questions, une nouvelle fenêtre **FIN DE LA SEANCE - VFERA** s'affiche.

– La première partie du questionnaire est terminée.

– Le patient peut tranquillement effectuer sa séance de thérapie.

B. A la fin de chaque séance :

- ⇒ **Vérifiez** que vous êtes bien sur la fenêtre intitulée **FIN DE LA SEANCE –VFERA**
Pour mémoire l’adresse de cette page est :
-



UNIVERSITÉ DE
BORDEAUX **FIN DE LA SEANCE - VFERA**

C'est un questionnaire en accès restreint. Vous devez avoir une invitation valide pour y participer.
Si vous avez reçu une invitation, saisissez-la dans le champ ci-dessous et cliquez sur Continuer.

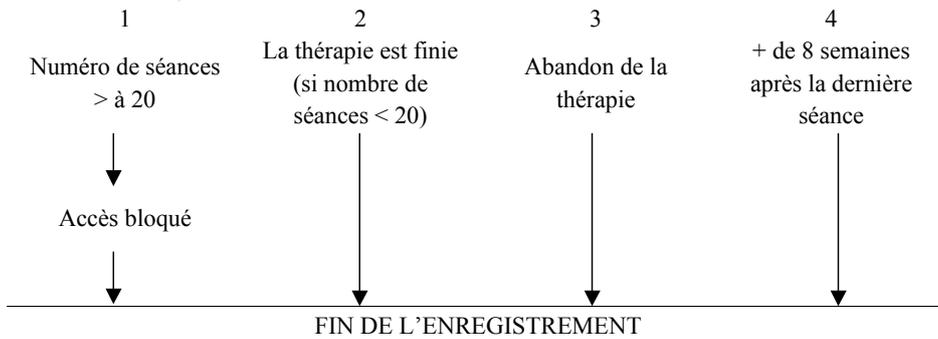
Code :

Continuer

- ⇒ **Entrez** de nouveau le code patient
- ⇒ **Cliquez** sur *continuer*
- ⇒ **Proposez** au patient de remplir seul les dernières questions
- Une fois que le patient a répondu à cette série de questions, l’écran revient automatiquement à la page **DEBUT DE LA SEANCE - VFERA** pour votre prochain patient.

C. Quand arrêter d'enregistrer les réponses des patients ?

- Le logiciel permet d'enregistrer les données jusqu'à 20 séances de thérapie et bloque ensuite l'accès au questionnaire.
- La fin de l'enregistrement des données par le thérapeute correspond à un de ces quatre cas de figure :



- Lorsque l'enregistrement des 3 patients est terminé (conditions ci-dessus) :
- ⇒ **Remplissez** les informations sur le thérapeute (zone 2) Informations sur le thérapeute (âge, sexe, profession, années d'expérience)
- ⇒ **Remplissez** les informations sur le patient (zone 3) Patients (patient mandaté, évaluation, approche thérapeutique utilisée, médication psychotrope, arrêt enregistrement).
- ⇒ **Renvoyez** la fiche à l'aide de l'enveloppe fournie affranchie.

université de BORDEAUX

ETUDE VFERA – FICHE DE POINTAGE
MERCİ D'ECRIRE EN LETTRES CAPİTALES

2 INFORMATIONS SUR LE THERAPEUTE

Age: ___ ans Sexe: 1. F 2. M Profession: 1. Psychologue 2. Psychiatre Année(s) d'expérience: ___ an(s)

PATIENTS					
CODE	Patient mandaté	Evaluation	Approche thérapeutique utilisée	Médication psychotrope	Arrêt enregistrement
	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/> 2. Trouble anxieux <input type="checkbox"/> 3. Dépendance et abus de substance <input type="checkbox"/> 4. Trouble psychotique <input type="checkbox"/> 5. Démence et trouble cognitif sévère <input type="checkbox"/> 6. Autre: _____ <input type="checkbox"/>	-----	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Nombre de séances n > 20 <input type="checkbox"/> 2. Fin de la thérapie n < 20 <input type="checkbox"/> 3. Abandon de la thérapie <input type="checkbox"/> 4. Délais > 8 semaines depuis la séance précédente <input type="checkbox"/>
	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/> 2. Trouble anxieux <input type="checkbox"/> 3. Dépendance et abus de substance <input type="checkbox"/> 4. Trouble psychotique <input type="checkbox"/> 5. Démence et trouble cognitif sévère <input type="checkbox"/> 6. Autre: _____ <input type="checkbox"/>	-----	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Nombre de séances n > 20 <input type="checkbox"/> 2. Fin de la thérapie n < 20 <input type="checkbox"/> 3. Abandon de la thérapie <input type="checkbox"/> 4. Délais > 8 semaines depuis la séance précédente <input type="checkbox"/>
	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/> 2. Trouble anxieux <input type="checkbox"/> 3. Dépendance et abus de substance <input type="checkbox"/> 4. Trouble psychotique <input type="checkbox"/> 5. Démence et trouble cognitif sévère <input type="checkbox"/> 6. Autre: _____ <input type="checkbox"/>	-----	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Nombre de séances n > 20 <input type="checkbox"/> 2. Fin de la thérapie n < 20 <input type="checkbox"/> 3. Abandon de la thérapie <input type="checkbox"/> 4. Délais > 8 semaines depuis la séance précédente <input type="checkbox"/>

Contact : Christophe CAZAUVIELH
 ☎ christophe.cazauviel@u-bordeaux.fr
 Laboratoire de Psychologie, Santé et Qualité
 de vie E44139 - 3^{er} étage de la Victoire -
 33076 Bordeaux Cedex

Aide au remplissage des colonnes *patients* :

- Le patient est-il mandaté ? Si le patient a entrepris de lui-même la démarche de venir vous consulter, cochez non. S'il est envoyé par un tiers en consultation, cochez oui
- Quelle(s) problématique(s) apparai(ssent) au cours du traitement ? Cocher la case correspondante. Si la case « *autre* » est cochée, précisez brièvement le motif des consultations.
- Quelle(s) approche(s) thérapeutique(s) utilisez-vous avec ce patient ?
- Le patient prend-il un traitement psychotrope ? Cocher la case correspondante.
- Arrêt enregistrement : cochez la case correspondante

Le système de recueil des informations permet d'anonymiser les données lors de leur enregistrement et de leur traitement.

V. En cas de problème, les instructions à donner au patient

CAS N°1

La fenêtre du navigateur a été fermée par erreur, ou, vous avez été déconnectés d'internet :

- ⇒ **1. Ouvrez** votre navigateur
- ⇒ **2. Cliquez** sur le lien dans le mail de participation **DEBUT DE LA SEANCE - VFERA** ou **FIN DE LA SEANCE-VFERA** selon qu'il s'agit du **DEBUT** ou de la **FIN** de séance
- ⇒ **3. Vérifiez** que s'affiche bien la fenêtre d'accès par code
- ⇒ **4. Entrez** le code patient
- ⇒ **5. Proposez** au patient de recommencer sa saisie

ATTENTION :

SI la fenêtre d'accès par code ne s'affiche pas (accès direct aux questions)

- ⇒ **Notez l'heure de l'erreur**
- ⇒ **Quittez votre navigateur pour le re-initialiser**
- ⇒ **Sur tablette/mobile fermez l'application complètement**
- ⇒ **Reprendre à partir de l'étape 1 ci-dessus**

Puis

- ⇒ **Envoyez un mail à ----- avec** Objet : « erreur »
et dans le texte du message précisez les 3 informations suivantes :
« Erreur navigateur »
La Date et l'heure d'accès
Le code du patient

CAS N°2

Le patient veut effacer ses réponses et recommencer :

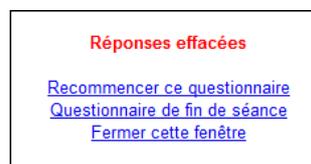
- ⇒ Cliquez en bas à gauche de l'écran sur le bouton *sortir et effacer vos réponses* (voir écran ci dessous).



- Un message de confirmation apparaît alors :



- ⇒ Cliquez sur *OK*.
- Une nouvelle fenêtre apparaît alors :



- ⇒ Cliquez sur *Recommencer ce questionnaire*

CAS N°3

Le patient veut effacer ses réponses et quitter définitivement le protocole de recherche :

⇒ Cliquez en bas à gauche de l'écran sur *sortir et effacer vos réponses*

- Un message de confirmation apparaît alors :

⇒ Cliquez sur OK.

- Une nouvelle fenêtre apparaît alors :

⇒ Cliquez sur *Fermer cette fenêtre*

OU

⇒ Fermez le navigateur.

⇒ Barrez ce patient sur la fiche de pointage et marquez **ABANDON RECHERCHE** comme ci-dessous :

université de BORDEAUX ETUDE VFERA – FICHE DE POINTAGE
MERCİ D'ECRIRE EN LETTRES CAPİTALES

INFORMATIONS SUR LE THERAPEUTE

Age: ___ ans Sexe: 1. F 2. M Profession: 1. Psychologue 2. Psychiatre Année(s) d'expérience: ___ an(s)

PATIENTS					
CODE	Patient mandaté	Evaluation	Approche thérapeutique utilisée	Médication psychotrope	Arrêt enregistrement
	1. Oui <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/>		1. Oui <input type="checkbox"/>	1. Nombre de séances n > 20 <input type="checkbox"/>
	2. Non <input type="checkbox"/>	2. Trouble anxieux <input type="checkbox"/>		2. Non <input type="checkbox"/>	2. Fin de la thérapie n = 20 <input type="checkbox"/>
		3. Trouble bipolaire <input type="checkbox"/>			3. Abandon de la thérapie <input type="checkbox"/>
		4. Trouble de la personnalité <input type="checkbox"/>			4. Délais > 8 semaines depuis la séance précédente <input type="checkbox"/>
		5. Démence et trouble cognitif sévère <input type="checkbox"/>			
		6. Autre : _____ <input type="checkbox"/>			

ABANDON RECHERCHE

Merci beaucoup pour votre implication et votre participation à cette recherche.

ANNEXE 9 : FICHE DE POINTAGE (VFERA)



ETUDE VFERA – FICHE DE POINTAGE
MERCI D'ECRIRE EN LETTRES CAPITALES

INFORMATIONS SUR LE THERAPEUTE					
Age	<input type="text"/> __ __ ans	Sexe	<input type="checkbox"/> 1. F <input type="checkbox"/> 2. M	Profession	<input type="checkbox"/> 1. Psychologue <input type="checkbox"/> 2. Psychiatre
				Année(s) d'expérience :	<input type="text"/> __ __ an(s)

PATIENTS					
CODE	Patient mandaté	Evaluation	Approche thérapeutique utilisée	Médication psychotrope	Arrêt enregistrement
	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/> 2. Trouble anxieux <input type="checkbox"/> 3. Dépendance et abus de substance <input type="checkbox"/> 4. Trouble psychotique <input type="checkbox"/> 5. Démence et trouble cognitif sévère <input type="checkbox"/> 6. Autre : _____ <input type="checkbox"/>	-----	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Nombre de séances n > 20 <input type="checkbox"/> 2. Fin de la thérapie n < 20 <input type="checkbox"/> 3. Abandon de la thérapie <input type="checkbox"/> 4. Délai > 8 semaines depuis la séance précédente <input type="checkbox"/>
	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/> 2. Trouble anxieux <input type="checkbox"/> 3. Dépendance et abus de substance <input type="checkbox"/> 4. Trouble psychotique <input type="checkbox"/> 5. Démence et trouble cognitif sévère <input type="checkbox"/> 6. Autre : _____ <input type="checkbox"/>	-----	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Nombre de séances n > 20 <input type="checkbox"/> 2. Fin de la thérapie n < 20 <input type="checkbox"/> 3. Abandon de la thérapie <input type="checkbox"/> 4. Délai > 8 semaines depuis la séance précédente <input type="checkbox"/>
	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Trouble de l'humeur <input type="checkbox"/> 2. Trouble anxieux <input type="checkbox"/> 3. Dépendance et abus de substance <input type="checkbox"/> 4. Trouble psychotique <input type="checkbox"/> 5. Démence et trouble cognitif sévère <input type="checkbox"/> 6. Autre : _____ <input type="checkbox"/>	-----	1. Oui <input type="checkbox"/> 2. Non <input type="checkbox"/>	1. Nombre de séances n > 20 <input type="checkbox"/> 2. Fin de la thérapie n < 20 <input type="checkbox"/> 3. Abandon de la thérapie <input type="checkbox"/> 4. Délai > 8 semaines depuis la séance précédente <input type="checkbox"/>

Contact : Christophe CAZAUVIEILH

Laboratoire de Psychologie, Santé et Qualité
 de vie EA4139 – 3 ter place de la Victoire –
 33076 Bordeaux Cedex

ETUDE VFERA - T

INSTRUCTIONS POUR LES ETUDIANTS

Contact : Christophe CAZAUVIEILH
Laboratoire de Psychologie, Santé et Qualité de vie EA4139
– 3 ter place de la Victoire –
33076 Bordeaux Cedex

Sommaire

I. PRE-REQUIS TECHNIQUES POUR PARTICIPER A L'ETUDE	3
II. ADMINISTRATION DES QUESTIONNAIRES	4
A. Première passation	4
B. Deuxième passation (15 jours plus tard)	5
III. EN CAS DE PROBLEME	6

I. Pré-requis techniques pour participer à l'étude

- Disposer d'une connexion internet.



- Utiliser un navigateur internet compatible.



L'application est compatible avec ces navigateurs et leurs versions ultérieures :

- SAFARI version 8.0.3
- FIREFOX version 33.1.1
- CHROME version 40.0.2214.115
- INTERNET EXPLORER version 11.0.9600.17691

! AVANT LA PREMIERE UTILISATION !

- ⇒ **Mettez à jour votre navigateur internet**
- ⇒ **Redémarrez votre ordinateur**

II. Administration des questionnaires

A. Première passation

⇒ **Connectez-vous** à internet.



⇒ **Ouvrez** le mail de participation à l'étude :
(ETUDE VFERA-T)



⇒ **Cliquez** sur le lien **VFERA – T**

Pour mémoire l'adresse est : -----

⇒ **Vérifiez** que s'affiche bien l'intitulé **VFERA - T**, dans la fenêtre du navigateur.
(Comme affiché ci-dessous)

UNIVERSITÉ DE
BORDEAUX **VFERA-T**

C'est un questionnaire en accès restreint. Vous devez avoir une invitation valide pour y participer.
Si vous avez reçu une invitation, saisissez-la dans le champ ci-dessous et cliquez sur Continuer.

Code :

⇒ **Notez la date du jour et Programmez-vous un rappel (alarme de téléphone ou calendrier) à 15 jours de cette date pour la seconde passation.**

⇒ **Rentrez le code** que vous avez reçu par mail.

⇒ **Cliquez** sur *continuer*

⇒ **Répondez** à toutes les questions

⇒ **Cliquez** sur *suivant* en bas de page.

- Une fois le questionnaire terminé, une nouvelle fenêtre apparaît :

Merci de votre participation.

⇒ **Fermez** la fenêtre de votre navigateur.

III. En cas de problème

CAS N°1

La fenêtre du navigateur a été fermée par erreur, ou, vous avez été déconnecté d'internet :

- ⇒ **1. Ouvrez** votre navigateur
- ⇒ **2. Cliquez** sur le lien dans le mail de participation **VFERA - T**
- ⇒ **3. Vérifiez** que s'affiche bien la fenêtre d'accès par code
- ⇒ **4. Entrez** votre code
- ⇒ **5. Recommencez** votre saisie

ATTENTION :

SI la fenêtre d'accès par code ne s'affiche pas (accès direct aux questions)

- ⇒ **Notez l'heure de l'erreur**
- ⇒ **Quittez votre navigateur pour le re-initialiser**
- ⇒ **Sur tablette/mobile fermez l'application complètement**
- ⇒ **Reprendre à partir de l'étape 1 ci-dessus**

Puis

- ⇒ **Envoyez un mail à _____ avec** Objet : « erreur »
et dans le texte du message précisez les 3 information suivantes :
« Erreur navigateur »
La Date et l'heure d'accès
Votre code d'accès

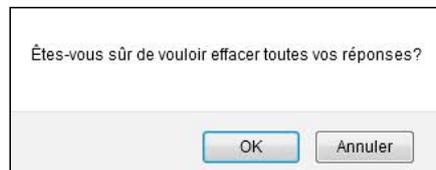
CAS N°2

Vous voulez effacer vos réponses et recommencer :

⇒ Cliquez en bas à gauche de l'écran sur le bouton *sortir et effacer vos réponses* (voir écran ci dessous).



- Un message de confirmation apparaît alors :



⇒ Cliquez sur *OK*.

- Une nouvelle fenêtre apparaît alors :



⇒ Cliquez sur *Recommencer ce questionnaire*

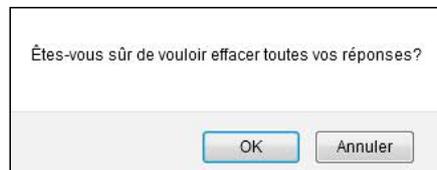
CAS N°3

Vous voulez effacer vos réponses et quitter définitivement le protocole de recherche :

⇒ Cliquez en bas à gauche de l'écran sur *sortir et effacer vos réponses*

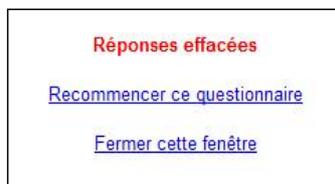


- Un message de confirmation apparaît alors :



⇒ Cliquez sur OK.

- Une nouvelle fenêtre apparaît alors :



⇒ Cliquez sur *Fermer cette fenêtre*

OU

⇒ Fermez le navigateur.

⇒ Envoyez un email au responsable de la recherche : -----

- avec pour objet « **ABANDON RECHERCHE** » et
- indiquez votre code dans le corps du message.

Merci beaucoup pour votre implication et votre participation à cette recherche.

NVB-O Study
Instructions for Therapists

TECHNICAL REQUIREMENTS

**PLEASE NOTE IT IS ESSENTIAL TO USE
FIREFOX BROWSER FOR THE TEST !!**

In order to participate, these following settings are REQUIRED :

- **The use of a computer :**
 - A desktop or a laptop
 - **NOT** a tablet (iPad...) because of the screen size
- **A good internet connection :** for playing the videos
- **A Firefox internet browser (free) :**
 - **Up to date**
 - Mandatory in order to ensure compatibility of specific questionnaires
 - Please download Firefox at the following address :
<https://www.mozilla.org/en-US/firefox/all/>
- **By default, Firefox enables JavaScript code execution :** you must not block javascript by using specific extensions.
- **You will need to activate your speakers volume on your computer** in order to hear the video clips sounds. You can also use headphones.

VERY IMPORTANT

You will need all of your attention during the survey. Please inactivate your mobile and computer alerts/messages, and do the experiment in a quiet room, where you can be alone.

If you have a problem/error loading a page during the survey, please click the arrow button of your Firefox browser  to re-load the page and to continue .

ACCESSING THE STUDY

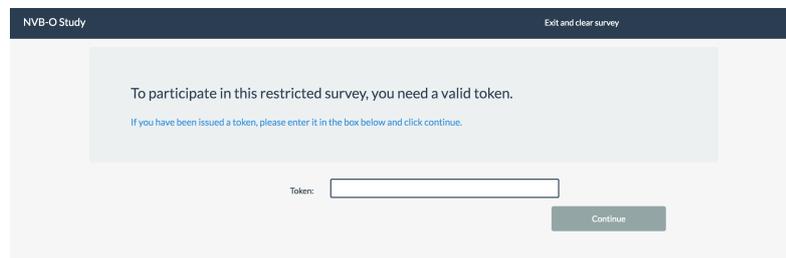
In the e-mail named « *NVB-O Study Access* », you have the following informations :

- **Your Token-Password in Red**
- The link to the **FIRST TEST**
- The link to the **RETEST** (14 days later)

FIRST TEST

REMEMBER YOU MUST USE FIREFOX AS YOUR BROWSER FOR THE TEST!!

- > Open your Firefox browser
- > Open the e-mail « *NVB-O Study Access* »
- > Copy/Paste the link **FIRST TEST IN THE FIREFOX** address bar and press Enter
(-----)
- > Enter your token/password in the screen displayed like below and click « *Continue* » :
- > Answer all the questions
- > After the last question a new page is displayed with the instruction « *This is a controlled survey. You need a valid token to participate* »
- > Enter your token/password in the screen displayed like below and click « *Continue* » :



- > Answer the questions
- > After the last question a new page is displayed with the title :
« *Test of Nonverbal Sensitivity - Short Form* »
DO NOT CLICK TO SWITCH LANGUAGE (THE TEST IS ALREADY IN ENGLISH)
- > Answer the series of questions (« Gender », « Age », « Are you student (currently) », « Do you use ») AND Enter your Token-password in the last text field named « Personal code ». **LEAVE THE LANGUAGE AS THE DEFAULT WHICH IS ENGLISH**
- > Click « *Continue* »
- > Answer the serie of questions
- > After the last question the following message is displayed : « *This is the end of this test. We thank you for your participation* »
- > Click « *Continue* », you are redirected to the page of the Swiss Center For Affective Science
- > Close Firefox

You have completed the FIRST TEST. RE-TEST WILL BE MUCH SHORTER (10 mn)

IMPORTANT : Please set a reminder in your agenda/calendar/phone, at 14 DAYS AFTER you filled this first questionnaire!!

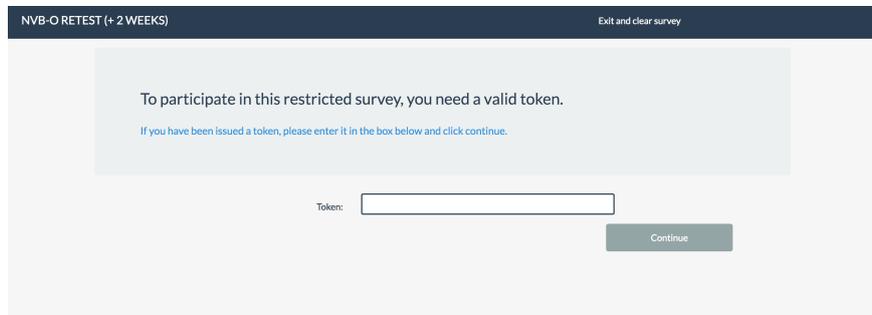
RETEST : 14 DAYS LATTER

REMEMBER YOU MUST USE FIREFOX AS YOUR BROWSER FOR THE TEST !!

- > Open your Firefox browser
- > Copy/Paste the following link into the Firefox browser address bar
<https://support.mozilla.org/en-US/kb/how-clear-firefox-cache>
- > Follow the tutorial to clear the temporary files « *Clear the cache* »

- > Open the e-mail « *NVB-O Study access* »
- > Copy/Paste the link RETEST INTO THE FIREFOX address bar and press Enter

- > Enter your token/password as displayed on the screen below :



NVB-O RETEST (+ 2 WEEKS) Exit and clear survey

To participate in this restricted survey, you need a valid token.
If you have been issued a token, please enter it in the box below and click continue.

Token:

Continue

- > Answer the series of questions
- > After the last question the following message is displayed : « *You have finished the FINAL TEST* »

The experiment is complete. Many thanks for your participation !

Christophe Cazauvielh

ANNEXE 12 : FORMULAIRE DE CONSENTEMENT ECLAIRE (NVB-O)

NVB-O Study	Exit and clear survey
<u>INFORMED CONSENT</u>	
<p>This research called "NVB-O Study" is conducted by Christophe Cazauviel, Psychologist.</p>	
<p>The different tasks in this survey are designed to assess your abilities in emotion recognition. This questionnaire will also ask about important elements of your professional status (education-training-supervision, type of practice) and personal abilities (the way you relate to your emotions, in you work).</p>	
<p>This research conforms with the french Psychologist deontological code :</p>	
<ul style="list-style-type: none">• The research supervisor guarantees that he will maintain the confidentiality of the research records or datas, and will ensure confidentiality in data treatments.• Datas will only be used for scientific research.• It is very unlikely that this research will cause you some distress, but in the case where you think you would need psychological support, please contact your general practitioner.• Participation in this study is voluntary and you are free to not participate. If you decide to participate, you may decide to stop participating at any time during the research, and ask the searcher to erase your data. You also have the right to access or rectify the records at the following adress :	
<p>Christophe CAZAUVIELH 5 rue du chai des Farines, 33000 Bordeaux/France contact@irpsy.com</p>	
<ul style="list-style-type: none">• The study is composed of the MIE-R task for emotion labelling, and of questions related to your profile (personal and professional). This material is hosted on the IRPPSY research website, and the data will be collected on this site. It also embodies the GERT-S survey for emotion recognition and the MINI-PONS Test for Nonverbal Sensitivity. The GERT- S is handled and hosted by the Swiss Center for Affective Sciences, Ghent University, and the European Consortium for Research on Emotional Competence Assessment (EURECA). The MINI-PONS is handled and hosted by the Swiss Center for Affective Sciences, Ghent University. By participating, you agree that your raw results (anonymized) on MINI-PONS and GERT-S will also be analysed by these institutions for research purposes.• By participating you agree to not copy, record, or disseminate, any of the different elements in this survey, by any processes. The video and questionnaires from this survey are protected by copyright, and the research supervisor received a specific authorization to use them in this research.	
<ul style="list-style-type: none">• You will have access to the general conclusions of this research, when the results will be available, if you want.	
<p>We thank you in advance for your participation in this research</p>	
<p>Your participation rest upon volunteering and requires the electronic signature (with time record) of this informed consent</p>	
<p><input type="radio"/> By checking this button, I specify I have been fully informed and I agree to participate in this research</p>	