

Validité empirique versus validité factorielle dans les inventaires de personnalité : le MMPI-2 et les échelles restructurées RC.

Empirical validity versus factorial validity in personality inventories: the MMPI-2 and the restructured RC scales

Michel Parisien ^a 

^aPsychologue

Abstract ■ Depuis 2003, plusieurs publications américaines ont paru sur les «Échelles Cliniques Restructurées RC». Ces échelles ont été construites en contradiction avec la stratégie psychométrique empirique adoptée d'emblée pour le MMPI (1943) et le MMPI-2 (1989). Les échelles RC sont plutôt basées sur une stratégie théorique-factorielle, que les experts de longue date des MMPI/MMPI-2 désavouent parce qu'elle ne remplit pas sa promesse de validité. En 2003, les échelles RC (Tellegen et coll., 2003) ont d'abord été ajoutées arbitrairement à l'ensemble des échelles empiriques du MMPI-2. Puis, en 2008 (voir Tellegen et Ben-Porath), elles ont plutôt été intégrées comme un atout à un tout nouvel instrument basé sur l'analyse factorielle, appelé MMPI-2-RF sur la seule base de l'emprunt de 338 des 567 items du MMPI-2. Le présent article a donc comme toile de fond la compromission possible du plus important test de personnalité utilisé en expertise psycholégale et en contexte clinique. Le but est de faire le point à la fois sur les fondements psychométriques du MMPI-2, ainsi que sur la valeur de ces échelles RC (Tellegen et coll., 2003) en affrontement maintenant ouvert avec le MMPI-2. ■ Since 2003, several American publications have appeared on the "Restructured Clinical Scales RC". These scales were constructed in contradiction with the empirical psychometric strategy adopted from the outset for the MMPI (1943) and MMPI-2 (1989). Rather, RC scales are based on a theoretical-factorial strategy, which longtime MMPI / MMPI-2 experts disavow because they do not fulfill their promise of validity. In 2003, these RC scales (Tellegen et al., 2003) were first arbitrarily grafted to the set of empirical MMPI-2 scales. Then, in 2008 (see Tellegen and Ben-Porath), they were instead integrated as an asset to a brand new instrument based on factor analysis, called MMPI-2-RF on the sole basis of the borrowing of 338 of the 567 items of the MMPI-2. The background to this article is therefore the possible compromise of the most important personality test used in forensic expertise and in a clinical context. The goal is to review both the psychometric foundations of MMPI-2, as well as the value of these RC scales (Tellegen et al., 2003) in the now open confrontation with MMPI-2.

Keywords ■ Empirical versus theoretical-factorial psychometric strategy; MMPI; MMPI-2; MMPI-2-RF; MMPI-3; Psychopathological clinical scales; Psychiatric diagnostic label versus identification marker; RC scales; MMPI-2 clinical scale covariation; Discriminant function analysis; Guttman facet theory.// Stratégie psychométrique empirique versus théorique-factorielle; MMPI, MMPI-2, MMPI-2-RF, MMPI-3; Échelle clinique psychopathologiques; Étiquette diagnostique psychiatrique versus identificatrice; Échelle RC; Covariation des échelles cliniques; Analyse discriminante; Théorie des facettes de Guttman..

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

 m.parisien.psy@mparisien.ca

 [10.20982/tqmp.17.2.p087](https://doi.org/10.20982/tqmp.17.2.p087)



Introduction

Parmi l'arsenal très volumineux des tests psychométriques et notamment les tests dits «inventaires de personnalité», le *Minnesota Multiphasic Personality Inventory* (1943), puis le MMPI-2, une version mise à jour en 1989, occupent l'un des premiers rangs grâce à une réputation établie dès la première moitié du XXe siècle et par sa crédibilité reconnue dans les domaines clinique et psycho-légal. Or, plus récemment, les auteurs des échelles restructurées *RC* (*MMPI-2 Restructured Clinical Scales*) proposées pour améliorer la version 2 du test taxaient les MMPI/MMPI-2 d'«empirisme aveugle» (TELLEGEN & coll., 2006). Ils tentaient ainsi de justifier leur propre approche théorique-factorielle (TELLEGEN & coll., 2003) pour de nouvelles échelles cliniques de base, les échelles *RC*, à intégrer au MMPI-2, approche dont la méthodologie de construction s'écarte de manière drastique des méthodes empiriques traditionnelles qui ont caractérisé dans leur essence même les inventaires MMPI et MMPI-2.

Pour situer le lecteur qui ne serait pas familier de ces inventaires, précisons que HATHAWAY et MCKINLEY (1940, 1942) commencèrent à construire le MMPI sous la bannière nouvelle de l'«empirisme» (voir Définitions/Précisions dans la section suivante). D'abord, le choix d'un bassin d'items ($N = 504$) se fit sans rationnel explicite particulier, et seule comptait au bout du compte la validation empirique finale de chaque item, basée sur des groupes-critères en contraste avec un échantillon normal, suite à l'échec relatif des inventaires de personnalité antécédents, dont la construction était basée sur le seul critère de validité manifeste de chaque item (approche essentiellement théorique et subjective), la table était mise pour cette position méthodologique, radicale pour l'époque, explicitée par MEEHL (1945b, 1945a). Ce dernier soulignait le danger d'accepter des items uniquement sur la base de leur contenu (voir également BERG, 1959), une position qui heurte de front le rationnel à la base des échelles *RC*. Le succès du MMPI fut fulgurant, donnant lieu (encore aujourd'hui) à du jamais vu en termes de recherches et d'utilisations cliniques (expertises légales, indications diagnostiques, orientations thérapeutiques). SELBOM et BEN-PORATH (2006) indiquent que les MMPI/MMPI-2 ont donné lieu à plus de 8 800 publications dans des revues scientifiques approuvées par des pairs, ce qui est exceptionnel dans ce domaine.

Hathaway et McKinley eurent recours également à des dispositifs subtils afin de détecter la validité du test en fonction des réponses données par les sujets. Rappelons l'intérêt d'Hathaway pour les détecteurs de mensonges (RUCHENNE, 2019).

Au lieu d'utiliser des ensembles indépendants de tests,

chacun avec un objectif particulier, HATHAWAY et MCKINLEY (1940) ont réuni dans un seul test un échantillon important de comportements pour examens par des psychologues, incluant un large éventail d'items, d'où émergent diverses échelles représentant une variété de descriptions valides de personnalité (telles que présentes à l'esprit clinique de cette époque). Malgré l'intention première de ces auteurs, l'étiquette «diagnostique psychiatrique» (voir Définitions/Précisions dans la section suivante) qu'ils ont voulu apposer aux échelles cliniques s'est révélée éphémère. On parlerait plutôt aujourd'hui d'étiquette «identificatrice». Ce sont les corrélats empiriques (des échelles individuelles, des paires d'échelles, des triades d'échelles) qui se sont avérés être comme une base solide de caractérisation de la personnalité et du comportement en référence à des numéros d'échelles, plutôt que de renvoyer à des désignations diagnostiques psychiatriques éventuellement périmées ou s'éloignant des nomenclatures psychiatriques sans cesse en évolution depuis le premier DSM (*Diagnostic and Statistical Manual of Mental Disorders* : voir AMERICAN PSYCHIATRIC ASSOCIATION, 1952).

Définitions/Précisions

Stratégie psychométrique empirique

Au MMPI, la stratégie psychométrique dite des «groupes contrastés» (comparaison de groupes) a permis à Hathaway (HATHAWAY & MCKINLEY, 1940, 1942, 1943), pour les échelles cliniques dites «psychopathologiques» (voir infra : *Les atouts du MMPI-2*) d'identifier les items discriminants spécifiques à chaque comparaison de sous-populations (p. ex. normaux versus dépressifs; normaux versus schizophrènes), et de réévaluer ces items par une validation croisée. Cette stratégie psychométrique est appelée «validité d'utilité pratique» par LAURENCELLE (1998), ou encore «mesure signifiante» (c.-à-d. permettant de discriminer à partir d'un critère) par CALDWELL (2006).

Ensuite, pour d'autres échelles des MMPI/MMPI-2 (comme la série des échelles de contenu), il s'agit de procédures rationnelles-empiriques, où une partie de l'élaboration inclut une validation par rapport à un critère.

Stratégie théorique-factorielle

Référant ici aux échelles cliniques restructurées *RC* (TELLEGEN & coll., 2003), Caldwell qualifie celles-ci de «mesure maximale» (c.-à-d. exprimant maximale la valeur), lesquelles pourtant ne résistent pas à l'épreuve de validité discriminante (diagnostique ou d'identification), comme on le verra dans ce qui suit. LAURENCELLE (1998) réfère quant à lui à la fragilité du concept de validité factorielle (tel qu'appliqué au 16PF de Cattell), basé essentiellement sur le contenu des libellés d'items.



Les échelles (ou indices) de validité

Le terme « validité » se rapporte ici au degré de confiance que peut avoir l'examineur sur le fait que les réponses du sujet sont représentatives de sa personnalité. Il renvoie donc à la « fiabilité » ou l'honnêteté objective des réponses du sujet et non pas sa seule « sincérité ». Certaines échelles (ou indices) de validité concernent la cohérence des réponses, d'autres les attitudes de minimisation de la psychopathologie, d'autres les attitudes d'exagération de la psychopathologie (voir PARISIEN, 2014).

Les échelles de contenu

L'interprétation de la série d'échelles appelées « de contenu » (voir BUTCHER, GRAHAM, WILLIAMS & BEN-PORATH, 1990; BEN-PORATH & SHERWOOD, 1993; GRAHAM, BEN-PORATH & McNULTY, 1999) se base sur le contenu des items et sur les corrélats empiriques présents dans la littérature. Leary (1957; voir Friedman et coll., 2015) a fait la distinction entre deux niveaux d'informations cliniques aux MMPI/MMPI-2, correspondant aux échelles cliniques versus les échelles de contenu. Le niveau I concerne la manière dont une personne est décrite par les autres ou la pression interpersonnelle que ses symptômes, plaintes, préoccupations, attitudes et traits de caractère exercent sur elle. Ces données sont publiques et objectives, et peuvent ou non être en accord avec la vision que la personne a d'elle-même ou de sa situation. C'est ce niveau qui se reflète dans le *profil clinique* du MMPI-2 ou dans les échelles dont la validité repose essentiellement sur un critère objectif, selon la stratégie empirique basée sur les groupes contrastés. Le niveau II concerne plutôt les descriptions conscientes du champ phénoménologique de la personne, le rapport de ses perceptions de soi en termes de comportement, de symptômes, de traits et de relations avec les autres. Les *échelles de contenu* sont particulièrement sensibles aux données relevant de ce niveau II. La stratégie rationnelle-théorique à la base des échelles RC est également basée sur le contenu des items.

MMPI versus MMPI-2

Rappel historique : stratégie psychométrique empirique.

Pour comprendre l'apport psychométrique essentiel du MMPI, le lecteur doit savoir que, dans les années 1910, la motivation majeure pour la construction de questionnaires de personnalité résidait dans le besoin d'évaluer l'aptitude du personnel militaire à servir dans les forces armées lors de la première guerre mondiale. Les questionnaires étaient construits sur une base rationnelle. Le

problème rencontré alors était le manque de contrôle vis-à-vis de l'attitude du sujet évalué, en particulier le penchant à suggérer ou exagérer une problématique personnelle dans le but d'éviter l'enrôlement.

Par ailleurs, la stratégie psychométrique des « groupes contrastés » (comparaison de groupes) a été proposée pour la première fois par Strong dans son « Inventaire d'intérêts vocationnels » (SVIB : STRONG, 1927). Strong avait bâti son inventaire en discriminant entre divers groupes « occupationnels », référant à leurs catégories d'emploi, ce dans la population masculine en général.

Pour le MMPI, cette méthode des « groupes contrastés » a plu au pragmatique Hathaway (voir RUCHENNE, 2019), qui se méfiait des théories, des préconceptions et des intuitions. Elle lui permit d'identifier les items discriminants spécifiques à chaque paire de sous-populations (p. ex. dépressifs versus normaux), puis de réévaluer ces items par une « validation croisée », les items ne répétant pas leur valeur discriminante significative étant mis de côté.

Le MMPI-2. Réécalonnage dans la continuité

Pour l'édition du MMPI-2 de 1989 (voir BUTCHER & coll., 2001), dans un esprit de continuité avec le MMPI original de 1943, il n'y a pas eu de re-validation des échelles cliniques déjà existantes, autant les échelles « de base » que les échelles « supplémentaires » (voir PARISIEN, 1999, 2014).

En outre, ont été mis de côté 106 des 566 items du MMPI. Du bassin original d'items, on a donc conservé 460 items, dont 68 ont été reformulés, et se sont ajoutés 89 nouveaux items pour de nouvelles échelles de contenu. Dix-huit des nouveaux items ont été rejetés, les vérifications empiriques menant à cette décision. Le MMPI-2 contient donc finalement 567 items.

Il y a eu des élagages minimes d'items dans les 13 échelles de base (échelles de validité et cliniques) ainsi que dans certaines échelles supplémentaires, de façon que les calculs effectués à partir des échelles originales et des nouvelles échelles produisent des scores à peu près équivalents.

En 1999, Parisien écrivait, à propos de la conférence de Butcher à Montréal le 18 mai 1989 :

Butcher a déjà affirmé (Parisien, 1989) que deux raisons majeures militaient en faveur d'une renormalisation du MMPI : 1. La vulnérabilité apparente de l'expert psycholégal au tribunal face à des normes vieilles de 50 ans;¹ 2. La non-comparabilité des résultats d'une échelle à l'autre à cause des distributions individuelles variables des cotes brutes (en termes d'asymétrie et de voussure) donnant lieu, selon les échelles, à des rangs-

1. De façon étonnante, les nouvelles normes du MMPI-2 se sont révélées assez proches de celles du MMPI.



centiles différents pour les mêmes cotes T linéaires.²

Les atouts du MMPI-2

NICHOLS (2011) nous indique que la force majeure du MMPI-2 réside dans sa continuité essentielle avec le MMPI. D'un instrument à l'autre, il n'y a que des changements mineurs, alors que les similarités sont considérables, tel que notamment, les stratégies empiriques du MMPI sont appliquées au MMPI-2.

Rappelons-nous que dans les années 1935-1940, les individus des groupes-critères du MMPI ayant servi à la construction des huit échelles cliniques dites psychopathologiques (voir plus loin tableau 1) n'ont pas été regroupés sur la base de diagnostics psychopathologiques précis (c'est une décennie plus tard qu'apparut le DSM-I; AMERICAN PSYCHIATRIC ASSOCIATION, 1952), mais plutôt sur la base de traits ou de symptômes communs tels que jugés par les auteurs du test. C'est pourquoi on préfère maintenant référer aux échelles cliniques par leur numéro plutôt que par leur nom (p. ex. échelle 8 plutôt qu'échelle «Schizophrénie»). En fait, pour le MMPI-2, essayer de constituer une toute nouvelle série d'échelles cliniques de base aurait présenté un risque considérable quant au maintien des liens laborieusement constitués entre les échelles cliniques d'origine et ces corrélats. Par ailleurs, malgré leurs imperfections, les propriétés des échelles cliniques de base sont très connues. Depuis les années 1940, le maintien des groupes-critères a permis l'accumulation de nombreux corrélats empiriques et les a comme protégées de l'interprétation première basée sur les concepts diagnostiques d'origine. Cette protection joue autant au MMPI-2 qu'au MMPI, puisque les nouvelles normes de l'échantillon normal du MMPI-2 de 1989 se sont montrées étonnamment proches de celles du MMPI de 1943 (voir GREENE, 1991; PARIEN, 1989).

NICHOLS (2011) énumère les avantages du MMPI-2 (avantages que ne présentent pas les échelles RC), dont :

- une stabilité temporelle des échelles cliniques de base, ni trop basse, ni trop élevée, reflétant à la fois les caractères de continuité et de changement des symptômes et de la personnalité chez un individu;
- une validité convergente (identification correcte des sujets dans les catégories réalisées) élevée, démontrée dans de très nombreuses recherches; et une validité prédictive modeste, mais solide;
- une capacité d'évaluer un très large éventail d'attitudes, de traits et de comportements, autant dans les populations normales que non normales;
- la disponibilité de plusieurs mesures qui, une fois

combinées, permettent une spécification relativement précise de l'attitude adoptée par le sujet pendant la période de testing et de la fiabilité de ses réponses;

- la disponibilité de plusieurs procédures interprétatives qui mettent l'accent soit sur les échelles individuelles, soit sur les patterns de profils, soit sur les contenus d'items, chacune offrant des contrôles et des élaborations potentielles sur les autres.

Mais le MMPI-2 a aussi des points faibles. En particulier, il y a d'importants chevauchements d'items entre plusieurs échelles cliniques, augmentant ainsi leurs intercorrélations et réduisant potentiellement leur validité discriminante. HELMES et REDDON (1993) imputent les problèmes de structure des MMPI/MMPI-2 à la redondance entre les échelles cliniques (et autres échelles), dont la cause est un chevauchement d'items qui gonfle les corrélations entre les différentes échelles. Ce problème de chevauchement d'items entre les échelles cliniques n'est pas anodin : dans les échantillons cliniques, l'intercorrélation moyenne entre les échelles cliniques de base se situe entre 0,55 et 0,60 (NICHOLS, 2011). HELMES et REDDON (1993) ont indiqué qu'il y a ainsi appauvrissement de la validité discriminante des échelles ainsi que de la structure factorielle du test. Bien sûr, la sensibilité de toutes ces échelles en est accrue, cela au détriment de leur spécificité. Ce chevauchement fut l'une des raisons majeures invoquées par les auteurs des échelles RC pour entreprendre leur tentative de remaniement majeur des échelles cliniques principales du MMPI-2.

Les échelles RC insérées au MMPI-2 en 2003 : brisure dans la stratégie psychométrique empirique

L'une des réformes les plus drastiques du MMPI-2 de 1989 vint de l'introduction en 2003 des échelles RC (Restructured Clinical Scales) dans le MMPI-2 Extended Score Report du service de correction informatisée NCS-Pearson Assessments : ces échelles RC, bien qu'étrangères sur le plan psychométrique à l'ensemble des MMPI/MMPI-2, faisaient désormais partie du protocole standard du MMPI-2. Dans les années 1990, Tellegen (avec d'autres auteurs du MMPI-2, dont Ben-Porath et Graham, mais sans Butcher) commença une recherche qui aboutit au début des années 2000 (TELLEGEN & coll., 2003) à une série de huit échelles cliniques restructurées RC, construites à partir d'une stratégie factorielle s'écartant franchement de la tradition empirique à la base des échelles cliniques des MMPI/MMPI-2. L'un des objectifs annoncés (mais non tenu) par le groupe de Tellegen était de contrer l'important chevauchement d'items entre les échelles cliniques. On commença donc par identifier les items communs aux

2. Il s'agit ici des huit échelles psychopathologiques 1, 2, 3, 4, 6, 7, 8 et 9. Afin de préserver la continuité avec les normes du MMPI, une transformation peu invasive des cotes brutes a été mise au point, via les cotes T uniformisées. Cette procédure est explicitée dans Greene (2000, 2011).



Tableau 1 ■ Correspondance entre les échelles RC et les échelles cliniques psychopathologiques

ÉCHELLES RC	ÉCHELLES CLINIQUES PSYCHOPATHOLOGIQUES
Facteur <i>dem</i> (échelle RCd : Démoralisation) des échelles RC	Facteur 1 (échelle A : Anxiété) des MMPI/MMPI-2
RC1/som (Plaintes somatiques)	1/Hs (Hypocondrie)
RC2/ipe (Peu d'émotions positives)	2/D (Dépression)
RC3/cyn (Cynisme)	3/Hy (Hystérie)
RC4/abs (Comportement antisocial)	4/Pd (Déviance psychopathique)
RC6/per (Idées de persécutions)	6/Pa (Paranoïa)
RC7/dne (Émotions négatives dysfonctionnelles)	7/Pt (Psychasthénie)
RC8/abx (Expériences aberrantes)	8/Sc (Schizophrénie)
RC9/hpm (Activation hypomaniaque)	9/Ma (Hypomanie)

Note. Si l'on ajoutait les 2 échelles non-psychopathologiques, soit 5(Mf ou Masculinité-féminité) et 0(Si ou Introversion sociale), on aurait là les 10 échelles du profil clinique de base des MMPI/MMPI-2.

échelles cliniques, items dont le regroupement a donné lieu à une mesure d'inadaptation générale (ou de détresse subjective) évoquant le facteur 1 du bassin d'items des MMPI/MMPI-2 (donc l'équivalent de l'échelle A, ou Anxiété, de WELSH, 1956). Ceci donna lieu au facteur *dem* du groupe d'échelles RC (voir tableau 1), soit *Demoralization* (RCd). L'identification des items RCd correspondit à l'assainissement des huit échelles cliniques psychopathologiques du MMPI-2 renommées par la suite RC1 à RC4, et RC6 à RC9, en purifiant celles-ci par l'identification (*capture*) du «noyau définitionnel distinctif» (*distinctive substantive core*) de chacune des huit anciennes échelles cliniques psychopathologiques (1, 2, 3, 4, 6, 7, 8, et 9) des MMPI/MMPI-2 (NICHOLS, 2011). Le tableau 1 présente la correspondance entre les échelles RC et les échelles cliniques MMPI/MMPI-2. La traduction française du nom des échelles RC est tirée de RUCHENNE (2019).

Il faut rappeler au lecteur que parmi les raisons du chevauchement d'items des échelles cliniques, il y en a deux qui sont majeures :

Le chevauchement de symptômes parmi les syndromes psychiatriques. Ce chevauchement augmente la sensibilité des échelles, tout en diminuant la spécificité de chacune (FRIEDMAN, GLESER, SMELTZER, WAKEFIELD & SCHWARTZ, 1983).

La variance commune avec le 1er facteur. Les analyses factorielles des échelles du MMPI, postérieures à la parution du test en 1943, ont régulièrement identifié deux facteurs qui ont été nommés et interprétés de diverses façons. WELSH (1956) a construit deux échelles pour représenter et mesurer ces deux facteurs, soit le facteur 1 (échelle A : Anxiété) et le facteur 2 (échelle R : Répression) (voir par exemple GREENE, 2011). Ce facteur 1, source principale de covariation entre les items des MMPI/MMPI-2, est associé à une inadaptation générale vaste et non-spécifique, ou à une dimension de détresse subjective. Il est diver-

sement représenté, par exemple à l'échelle A (Anxiété : WELSH, 1956), ou encore au facteur RCd/*dem* (*Demoralization* : TELLEGEN & coll., 2003). C'est un facteur envahissant parmi l'ensemble des items du test et il gonfle les intercorrélations entre la plupart des échelles en compromettant possiblement leur validité discriminante.

C'est par référence à ce contexte que le projet des échelles restructurées RC a été entrepris dans le dessein de créer une série d'échelles qui reflètent mieux les «construits conceptuellement significatifs et cliniquement importants» (TELLEGEN & coll., 2003) présents dans les échelles cliniques d'origine. Ainsi, «les échelles RC ont été conçues pour préserver les propriétés descriptives importantes des échelles cliniques existantes du MMPI-2, en améliorant leur caractère distinctif» (ibid.).

BUTCHER (2011) a critiqué la stratégie de création des échelles RC, en particulier l'abandon par TELLEGEN & coll. (2003) de la méthode empirique de construction sur laquelle reposait les MMPI/MMPI-2. À cet égard, CALDWELL (2006) a noté que la différence entre les échelles extraites par analyse factorielle (les échelles RC) et les échelles établies empiriquement par la méthode des groupes contrastés versus la population normale (impliquant les échelles cliniques psychopathologiques des MMPI/MMPI-2) réside dans ce qu'il nomme «mesure maximale» (c.-à-d. pour mesurer très bien quelque chose) versus «mesure significative» (c.-à-d. pour discriminer à partir d'un critère). De façon similaire, LAURENCELLE (1998, chapitre 3) mentionne les tests 16PF et MMPI pour illustrer respectivement la «validité théorique» versus l'«utilité pratique» d'un test, la fragilité du concept théorique de validité factorielle (telle que démontrée traditionnellement et correspondant au 16PF de Cattell) étant actuellement mise en cause (ANDRÉ, LOYE & LAURENCELLE, 2015). La naïveté des concepts issus de l'analyse factorielle classique tient principalement au fait que ces concepts soient basés sur une lecture et une



interprétation intéressée des items par le répondant c'est-à-dire sa façon de se voir lui-même, l'image qu'il veut projeter aux autres. C'est un vice intrinsèque des tests auto-rapportés. Dans les années 1940, Hathaway (voir HATHAWAY & MCKINLEY, 1943) en avait une conscience aigüe (voir RUCHENNE, 2019), d'où son recours aux stratégies empiriques décrites dans le présent article ainsi que l'introduction d'indices et échelles de validité (fiabilité) des réponses.

Origines des échelles RC

Il est ici essentiel de comprendre les origines et le développement des échelles restructurées RC, car nous entendons démontrer que, dès le départ, les auteurs TELLEGEN & coll. (2003) ont fait fausse route. En 1985, Watson et Tellegen cherchaient à définir un «concept d'humeur» qui fasse consensus, et ils ont préconisé la théorie de deux systèmes d'activation générale de l'affect, représentés par la configuration dimensionnelle PA versus NA (affect positif versus affect négatif : voir WATSON & TELLEGEN, 1985; TELLEGEN, WATSON & CLARK, 1999; D. WATSON, WIESE, VAIDYA & TELLEGEN, 1999), configuration fondée sur un modèle factoriel de la structure affective et appliquée à l'humeur auto-évaluée. Ils ont développé un modèle dit «hiérarchique» à trois niveaux pour décrire la structure de l'effet ressenti.

En 2003, les auteurs des échelles restructurées RC (TELLEGEN & coll., 2003) ont utilisé le cadre du modèle Watson-Tellegen pour façonner et clarifier le concept de «démoralisation» (c.-à-d. la base de l'échelle RCd/dem) visant à cerner la source de la covariation problématique entre les échelles cliniques du MMPI-2. Pour les auteurs, la dimension PU (*Pleasant* versus *Unpleasant*), qui se situe au sommet de leur hiérarchie, témoignerait de cette covariation. Qui plus est, cette dimension PU serait, selon eux, omniprésente dans d'autres inventaires cliniques auto-rapportés. C'est en se fondant sur cette hypothèse que les auteurs des échelles RC ont extrait la dimension de démoralisation des échelles cliniques du MMPI-2 et l'ont confinée à leur propre échelle RCd afin de créer, par la suite, un ensemble de huit échelles RC qui soient davantage conceptuellement distinctes.

Le clinicien non informé pourrait supposer que le modèle hiérarchique de Watson et Tellegen est universellement accepté et appliqué là où un modèle de structure d'affect est nécessaire. Or, il n'en est rien. Il existe un certain nombre de modèles concurrents (LARSEN & DIENER, 1992; RUSSELL, 1980; THAYER, 1996), largement validés et plus importants que le modèle hiérarchique (p. ex. le modèle *circumplex* de Russell), et qui décrivent aussi la structure de l'affect. Le modèle de Watson et Tellegen a été vigoureusement remis en cause depuis 1985. À la suite de plusieurs études et débats à la fin des années 1990, de nom-

breuses lacunes du modèle hiérarchique ont été mises en évidence (voir CARROLL, YIK, RUSSELL & FELDMAN-BARRETT, 1999; GREEN, SALOVEY & TRUAX, 1999; RANSON, NICHOLS, ROUSE & HARRINGTON, 2009).

Développement des échelles RC

Dès le début, l'allégeance des auteurs à leur théorie hiérarchique semble avoir étouffé leur curiosité pour des modes de résolution plus simples afin de réduire l'intercorrélation entre les échelles cliniques psychopathologiques du MMPI-2 ou encore pour définir la base de la covariance du facteur 1, à partir de repères validés empiriquement. Leur méthode de construction aurait donc été biaisée en faveur d'une approche inutilement dogmatique, excluant l'exploration de méthodes empiriques qui auraient maintenu la tradition des MMPI/MMPI-2.

Dans le chapitre 3, *Developing the Restructured Clinical Scales* du manuel édité par *Pearson Assessments* (TELLEGEN & coll., 2003), les auteurs décrivent les étapes spécifiques suivies dans la construction des échelles RC, ce qui devrait éclairer le lecteur sur les bases méthodologiques de la création des échelles RC.

Étape 1 : échelle RCd

Les auteurs ont identifié un sous-ensemble d'items pour mesurer le 1er facteur et créer une nouvelle échelle, appelée *Demoralization* pour représenter ces items. Ce processus s'est inspiré du modèle hiérarchique de Watson-Tellegen, selon lequel la variance du 1er facteur des MMPI/MMPI-2 correspondait à la dimension PU (*Pleasant* versus *Unpleasant*) de ce modèle.

Pour créer les items de l'échelle *Demoralization* (*dem* : $N = 23$ items), les auteurs ont d'abord combiné les items des échelles qu'ils ont jugées les plus saturées de la dimension PU, soit les échelles 2 (Dépression) et 7 (Psychasthénie). Puis, ils ont effectué diverses analyses corrélationnelles afin de réduire à l'essentiel le contenu de l'échelle, aboutissant finalement à un total de 10 items. Treize autres items ont été tirés du reste du bassin sur la base de leurs corrélations avec les facteurs du modèle Watson-Tellegen, soit PEM (*Positive Emotionality*) et NEM (*Negative Emotionality*).

Étape 2 : composants de base

Les auteurs ont tenté d'évacuer de chacune des échelles cliniques psychopathologiques du MMPI-2 la covariance qu'elle avait avec le facteur *dem*, toujours par des techniques corrélationnelles et factorielles sophistiquées. Ils ont ainsi défini une dimension jugée refléter un «noyau définitionnel» pour chaque échelle et qui reste distincte de l'échelle *Demoralization* (tout en lui restant corrélée) ainsi que des composantes essentielles identifiées dans d'autres



échelles cliniques du MMPI-2.

Étape 3 : échelles de départ

Les auteurs ont sélectionné 158 items des échelles cliniques d'origine comme matière à former de nouvelles versions d'échelles, items qui ont ensuite été sélectionnés pour réduire le chevauchement inter-échelles et augmenter la cohérence intra-échelles. Sur la base des items restants, un second ensemble d'échelles a été formé auquel lesdits items ont été ajoutés ou supprimés afin d'augmenter le caractère distinctif (associé au «noyau définitionnel») de chaque échelle. Les 73 items ayant survécu à ces procédures ont ensuite été triés dans un ensemble final d'échelles de départ, ceci pour aboutir aux échelles *RC1*, *RC2*, *RC3*, *RC4*, *RC6*, *RC7*, *RC8* et *RC9*.

Étape 4 : échelles RC finales

Les échelles de départ ont été augmentées enfin par des items tirés du bassin d'items MMPI-2. En bref, un item était ajouté à une échelle de départ s'il démontrait une bonne convergence avec cette échelle en même temps qu'une faible corrélation avec les autres échelles : ces procédures ont abouti aux échelles *RC* finales.

Critiques des échelles RC

Toujours dans la perspective de jauger la valeur et l'utilité pratique des échelles restructurées *RC*, issues de stratégies psychométriques factorielles, nous abordons maintenant quatre groupes de critiques amenées par différents experts du MMPI-2, avec comme considération commune la valeur supérieure des échelles cliniques empiriques des MMPI/MMPI-2.

Critique 1 : mesure «maximale» versus mesure «signifiante». Nous reprendrons d'abord ici ce que CALDWELL (2006) amenait comme critiques de fond envers les échelles restructurées *RC* dans le volume 87(2) du *Journal of Psychological Assessment* qui consacrait 11 articles aux échelles restructurées *RC* du MMPI-2, dont l'article de Caldwell.

1. Caldwell mentionne que la tâche d'extraire des huit échelles cliniques psychopathologiques (1, 2, 3, 4, 6, 7, 8, et 9) l'aspect Démoralisation (soit d'extraire le facteur 1, ou encore l'échelle A) n'a pas bien fonctionné, en regard des corrélations persistantes entre les échelles *RC* et des mesures indépendantes représentant ce facteur 1;
2. il écrit que le but d'aboutir à une «mesure parfaite» (*maximal measurement*) ne paraît pas réaliste;
3. selon lui, cet essai d'extraction tenté par les auteurs des échelles *RC* dénature la complexité et la richesse des échelles cliniques de base des MMPI/MMPI-2, alors que les échelles empiriques validées sur des groupes-critères, comme celles des MMPI/MMPI-2, aboutissent

à une «mesure signifiante» (*meaningful measurement*) de caractéristiques individuelles qui se rapprochent du groupe-critère dont ces échelles émanent;

4. l'interprétation des échelles *RC* est uniquement basée sur la «validité manifeste» (*face validity*) des items qui les composent. Or, le postulat que la perception de soi et l'autodescription aboutissent habituellement à des résultats valides peut être trompeur, voire fallacieux, eu égard aux objectifs d'évaluation et d'intervention;
5. il y a un contraste évident entre l'étroitesse de la spécificité interprétative des échelles *RC* et l'étendue (ainsi que la surprenante spécificité) caractéristiques des recherches concernant les codes-types du MMPI (HATHAWAY & MCKINLEY, 1943) et du MMPI-2 (BUTCHER, DAHLSTROM, GRAHAM, TELLEGEN & KAEMMER, 1989). Au MMPI-2, un code-type correspond aux chiffres d'identification des deux échelles dont la cote (ou score) *T* (de moyenne 50 et écart-type 10) excède 65 au profil clinique de base. Le 1er chiffre représente l'échelle la plus élevée, chaque échelle étant codée par un chiffre (voir tableau 1). Par exemple, le code-type 6-3 représente les échelles **Pa** et **Hy**. Ce sont précisément ces différents codes-types qui ont été par la suite empiriquement établis, permettant l'accumulation de nombreux corrélats comportementaux.

Critique 2 : allégeance à une approche théorique plutôt qu'empirique. Les problèmes minant les échelles *RC* proviennent de la première étape, soit le développement de l'échelle *RCd* *Demoralization*, pour laquelle les auteurs ont cherché à construire une «échelle-témoin» (en anglais, *marker*, ou «marqueur») mesurant «la large dimension émotionnellement colorée, capturée par exemple par l'échelle A bien connue de WELSH (1956, 2000), parfois appelée «1er facteur MMPI-2» (TELLEGEN & coll., 2003). Refusant l'empirisme qui guidait les pionniers du MMPI, il leur fallait un contexte théorique applicable à la construction de ce «marqueur», qu'ils ont interprété comme équivalent de leur axe *Pleasantness* versus *Unpleasantness* du modèle hiérarchique de l'affect positif versus négatif de WATSON et TELLEGEN (1985), qu'ils ont appelé *Demoralization* (*dem*).

Il y avait d'autres options de conceptualisation d'un «marqueur» empiriquement modifiées et disponibles en remplacement de ce facteur *Demoralization*. Ces options sont élaborées dans RANSON & coll. (2009) :

1. l'échelle A (Anxiété) de WELSH (1956);
2. des analyses factorielles plus récentes faites sur l'ensemble des items du MMPI-2 et identifiant un facteur : d'abord JOHNSON, NULL, BUTCHER et JOHNSON (1984), avec un facteur qu'ils ont nommé JB1, les lettres JB représentant deux auteurs de l'article cité ici, soit Johnson et Butcher; ensuite WALLER (1999), avec un facteur appelé W1, la lettre W représentant l'auteur Waller;



3. une option plus qu'évidente et à la disposition des auteurs des échelles *RC* (TELLEGEN & coll., 2003) aurait été de factoriser leurs propres données.

Aucune de ces options ne semble avoir intéressé ces auteurs des échelles *RC*. Il semble donc que c'est l'allégeance résolue à une approche théorique (plutôt qu'empirique) qui a guidé ces auteurs. Cette approche théorique leur a imposé un biais qui les a finalement éloignés de leurs objectifs initiaux.

D'abord, les procédures pour construire le concept/échelle *dem* ont ciblé et retenu des items qui ne se trouvent sur aucune des échelles cliniques du MMPI-2. En effet, ces items hors échelle représentent près de la moitié de la taille de l'échelle *dem* (voir RANSON & coll., 2009). Ainsi, l'aspiration à cerner la source la plus pertinente de la covariation entre les échelles cliniques du MMPI-2 a été abandonnée.

Ensuite, la décision d'adopter le concept théorique de démoralisation a introduit un écart par rapport à la tradition empirique des MMPI/MMPI-2, de la même façon que l'a fait l'adoption d'une procédure de recherche du concept de démoralisation qui a exercé une influence biaisante sur ce qui a émergé comme noyau fondamental distinctif de chaque échelle clinique au fur et à mesure que les procédures de restructuration progressaient.

Critique 3 : questionnements recommandés en 1978, mais évités en 2003. Des règles d'inclusion de nouvelles échelles au MMPI avaient déjà été émises une dizaine d'années avant la parution du MMPI-2 (voir BUTCHER & TELLEGEN, 1978).³ On s'accordait alors sur le fait que seules les nouvelles échelles qui produisent des résultats plus fidèles et plus valides, ou encore celles évaluant des concepts non-mesurés par les échelles existantes, satisfaisaient aux critères d'inclusion. Plusieurs chercheurs (voir par exemple BUTCHER, GRAHAM & BEN-PORATH, 1995) ont indiqué l'importance de comparer les nouvelles échelles MMPI-2/MMPI-A⁴ à toutes les échelles existantes acceptées et servant déjà de normes. Or, l'inclusion des échelles *RC* au MMPI-2 en 2003 suscitait des questionnements, en particulier par rapport aux aspects suivants : redondance, dérive conceptuelle et sensibilité.

Redondance

On a constaté (ROUSE, GREENE, BUTCHER, NICHOLS & WILLIAMS, 2008) que chacune des échelles *RC* est plus fortement corrélée au MMPI-2 soit avec une échelle supplémentaire, soit avec une échelle de contenu, soit avec l'une des échelles *PSY-5*,⁵ qu'avec son échelle clinique parente. Voilà une première redondance. Et pourtant, un grand nombre d'échelles *RC* (*RCd*, *RC1*, *RC3*, *RC7*, *RC8* et *RC9*) demeurent également redondantes avec les échelles cliniques MMPI-2 existantes en raison des corrélations relativement élevées qu'elles montrent avec ces échelles (mais pas autant élevées qu'avec certaines échelles de contenu ou autres du MMPI-2). Nonobstant les efforts de justification de TELLEGEN, BEN-PORATH et SELBOM (2009), GREENE, ROUSE, BUTCHER, NICHOLS et WILLIAMS (2009) ont présenté une réplique à ces auteurs dans laquelle ils ont, une fois de plus, mis en lumière les corrélations relativement élevées de plusieurs échelles *RC* avec les échelles MMPI-2 existantes, notant que le manuel technique MMPI-2-RF (TELLEGEN & BEN-PORATH, 2011) montre également les mêmes corrélations élevées que celles notées par ROUSE & coll. (2008).⁶ Cela a pour effet que les échelles *RC* sont présentées par leurs auteurs comme plus nouvelles qu'elles ne le sont en réalité.

On constate par ailleurs que les corrélations (*r*) les plus élevées⁷ entre les échelles *RC* et celles du MMPI-2 émanent dans ce dernier instrument des échelles d'items à contenu évident : soit la série «échelles de contenu», soit certaines échelles supplémentaires, soit des échelles *PSY-5* issues d'analyses factorielles, soit encore une échelle clinique particulière sans items à contenu subtil (échelle *Hs* non-corrigée par *K*; pour des informations sur l'échelle *K* et la correction-*K*, voir ci-après, section *Sensibilité*).

Dérive conceptuelle

NICHOLS (2006b) utilise le terme «dérive conceptuelle» (*construct drift*) pour faire référence à la possibilité que l'ajout d'items en corrélation avec des items d'amorçage (c.-à-d. l'étape 4 de la création des échelles *RC* : voir plus haut *Développement des échelles RC*) encourt une dérive conceptuelle en contradiction avec le noyau de départ. Ici sont concernées les échelles *RC3*, *RC4*, *RC7* et *RC9*.

3. Ce second auteur, Tellegen, s'avérera d'abord le co-auteur du MMPI-2 (BUTCHER & coll., 1989), et plus tard l'auteur principal des échelles restructurées *RC* (TELLEGEN & coll., 2003).

4. Le MMPI-A a été mis au point pour les jeunes de 14 à 18 ans (BUTCHER & coll., 1992).

5. Les échelles *PSY-5* (*Personality Psychopathology Five*) représentent une approche dimensionnelle basée sur une stratégie factorielle (HARKNESS, McNULTY & BEN-PORATH, 1995). Elles se nomment : 1. Agressivité; 2. Psychoticisme; 3. Désinhibition; 4. Neuroticisme/Névrosisme; 5. Introversion/Peu de réactions émotionnelles positives.

6. Le MMPI-2-RF (TELLEGEN & BEN-PORATH, 2011) est une forme dite restructurée du MMPI-2, soit une mesure d'auto-évaluation de 338 éléments comprenant 51 échelles de validité et de substance clinique nouvelles et révisées, avec les échelles cliniques restructurées (*RC*) comme fondement. Il a été développé avec le soutien de l'*University of Minnesota Press*, détenteur des droits d'auteur du MMPI-2, à la suite de la publication des échelles *RC* (2003).

7. Ces coefficients *r* se retrouvent soit dans GREENE (2011), soit dans FRIEDMAN, BOLINSKEY, LEVAK et NICHOLS (2015).



D'abord (voir BUTCHER, HAMILTON, ROUSE & CUMELA, 2006; NICHOLS, 2006b), l'échelle RC3 aurait été essentiellement (et volontairement) modifiée au point de devenir une échelle totalement différente de l'échelle clinique psychopathologique **Hy(3)** du MMPI-2 (voir tableau 1 dans le présent texte). BUTCHER (2011) a noté que les riches descripteurs ainsi que les codes-types associés à l'échelle clinique **3(Hy)** sont maintenant devenus obsolètes avec l'échelle RC3, qui elle-même est devenue redondante avec les échelles ou sous-échelles du MMPI-2 CYN/CYN1 (CYN = échelle de contenu «Cynisme»; CYN1 = sous-échelle de contenu «Croyances misanthropiques»). THOMAS et YOUNG-JOHN (2009) ont en outre noté que l'échelle RC3 a perdu toute utilité comme repère de somatisation chez les patients traumatisés cranio-cérébraux. En somme, à l'échelle RC3, en comparaison avec l'échelle **Hy(3)**, la mesure de l'expression des symptômes somatiques disparaît tout à fait, en même temps que les éléments de «démoralisation». L'échelle RC3 est devenue une échelle opposée au concept de naïveté typique de l'hystérie, pour être plutôt identifiée au cynisme (GORDON, 2006).

Ensuite, l'échelle RC4 a été l'objet de préoccupations (NICHOLS, 2006b; BOLINSKEY & NICHOLS, 2011) parce que la forte proportion d'items relatifs à la toxicomanie qui la composent présente un risque d'inférences faussement positives par rapport à des dispositions et comportements antisociaux.

Enfin, BOLINSKEY et NICHOLS (2011) ont affirmé que les échelles RC7 et RC9 ne mesurent pas les mêmes traits et comportements que les échelles d'origine. Ils proposent aux cliniciens des précautions à prendre quant aux interprétations suggérées par les auteurs des échelles RC.

Sensibilité

Des échelles homogènes avec uniquement des items au contenu manifeste ou évident, c'est-à-dire reconnaissable quant à son lien psychopathologique ou problématique, par exemple la série des échelles de contenu du MMPI-2 ainsi que les échelles restructurées RC, sont plus vulnérables aux attitudes de «minimisation» ou d'«exagération» psychopathologiques, que ne le sont des échelles cliniques multivariées plus complexes, par exemple la grande majorité des échelles cliniques psychopathologiques du MMPI-2. La présence dans une échelle clinique à la fois d'«items évidents» et d'«items subtils» (*obvious* versus *subtle*) constitue un rempart relatif, une précaution contre les sujets malhonnêtes ou malavisés présentant volontairement ou non un faux portrait d'eux-mêmes. C'est pourquoi des études utilisant une variété

d'échantillons ont révélé que les échelles RC montrent des niveaux de scores moins élevés que les échelles cliniques du MMPI-2, étant plus sujettes aux faux négatifs (associés aux attitudes de «minimisation» ou au manque d'introspection).

Il faut noter que ce sont les auteurs des sous-échelles cliniques Wiener et Harmon qui ont identifiés 146 items ouverts et 110 items subtils pour cinq (5) des échelles cliniques, soit : **2(D)**, **3(Hy)**, **4(Pd)**, **6(Pa)**, et **9(Ma)** (WIENER, 1948). Les trois (3) autres échelles, soit **1(Hs)**, **7(Pt)**, et **8(Sc)**, ne comportent pas d'items subtils, ce qui fait de ces échelles des candidates privilégiées pour les attitudes de «minimisation» ou d'«exagération» (voir GREENE, 1991, 2000, 2011). Ce n'est donc pas par hasard que dans le profil corrigé par K, un score de crédibilité des réponses, ce sont ces trois échelles qui bénéficient de la plus importante addition de cotes brutes **K** : **Hs.5K**, **Pt1K**, et **Sc1K**. Au MMPI, la correction-K (conservée au MMPI-2) correspond à une proportion de cotes brutes **K** ajoutée à certaines échelles cliniques pour compenser l'effet de simulation, variable selon les données empiriques obtenues à chaque échelle.⁸ Les normes fournissent également une option de tracé de profil «sans correction-K».

Par ailleurs, WALLACE et LILJEQUIST (2005) ont rapporté que les cotes *T* moyennes des patients cliniques sur les échelles RC étaient significativement inférieures aux cotes *T* des échelles cliniques d'origine du MMPI-2. La majorité des profils RC (56%) présentaient moins d'élévations qu'aux échelles cliniques d'origine (cf. MMPI/MMPI-2).

Dans un large échantillon (> 2000) de criminels incarcérés, MEGARGEE (2006) a constaté que les scores moyens aux échelles RC étaient généralement inférieurs aux scores RC moyens de l'échantillon normal, en fait en deçà d'une cote *T* de 56.⁹ Cela advient aussi pour RC4 (comportement antisocial), où l'on s'attendrait à un résultat moyen significativement élevé chez des détenus.

PIZITZ et McCULLAUGH (2011), dans un échantillon d'hommes harceleurs condamnés, indiquent que cinq échelles RC (RC2, RC3, RC7, RC8 et RC9) présentaient un score moyen inférieur à une cote *T* de 50, la moyenne populationnelle. Dans le cas de RC4 (comportement antisocial), une échelle dont on s'attendrait à ce qu'elle soit élevée dans une telle population, le score moyen n'était que de 51,7, soit inférieur à plus qu'un écart-type au score obtenu par ces hommes sur l'échelle clinique **4(Pd)** du MMPI-2.

Dans une étude sur la garde familiale (ARCHER, HAGAN, MASON, HANDEL & ARCHER, 2012), échantillon de sujets *a priori* non cliniques, mais naturellement enclins à la minimisation, la plupart des moyennes aux échelles RC étaient

8. Cette valeur proportionnelle au score de l'échelle **K** est simplement ajoutée au score de l'échelle indiquée, p. ex. le score brut X_{H_s} de l'échelle **Hs** devient $X_{H_s} + 0,5 \times X_K$.

9. Une cote (ou un score) *T* réfère à une distribution statistique de résultats bruts d'une échelle ou d'un indice où la moyenne est de 50 et l'écart-type de 10.



inférieures à la norme T de 50, à une exception près : l'échelle $RC6$ obtenant une cote T moyenne de 51,16.

Par ailleurs, GORDON, STOFFEY et PERKINS (2013) ont mené une étude empirique comparant la sensibilité des échelles RC dans la détection de problèmes de santé mentale dans un échantillon de patients cliniques. Les résultats ont révélé que la plupart des échelles cliniques du MMPI-2 ont une sensibilité plus élevée que les échelles RC à tous les niveaux de psychopathologie, et en particulier aux niveaux moins pathologiques. Les auteurs ont aussi déterminé que les différences obtenues n'étaient pas imputables à l'utilisation de la correction- K du MMPI-2.

Saborio et Hass (2012; voir Butcher et coll., 2015) ont mené une évaluation empirique à l'aide du MMPI-2 sur 167 femmes abusées sexuellement. Ils ont trouvé des différences très importantes entre les échelles cliniques du MMPI-2 et les échelles RC , celles-ci restant toutes à un niveau normal, à l'exception du facteur de démoralisation RCd ($T = 65$). Pourtant, ces victimes présentaient de graves problèmes de santé mentale : symptômes et comportements concordant avec les informations apportées lors des entretiens cliniques. Les auteurs confirment que les résultats aux échelles cliniques du MMPI-2 reflétaient par ailleurs les problèmes psychologiques rapportés.

Un exemple plus médiatisé concerne le tueur en série Theodore Kaczynski, surnommé *Unabomber*, docteur (Ph.D.) en mathématiques de l'Université du Michigan en 1967. Il fut professeur à l'Université de Californie pendant deux ans avant de mener une vie de reclus en forêt au Montana, entretenant des croyances anti-gouvernementales et antitechnologies. Pendant un certain nombre d'années, il envoya par la poste des bombes artisanales, soit à des universités, soit à American Airlines. Ainsi, il tua 3 personnes et en blessa 23. En 1998, il refusa d'admettre qu'il avait des problèmes mentaux. Pourtant, il tenta de se suicider durant les procédures légales. Il fut trouvé coupable de meurtre, et écopa d'une sentence d'emprisonnement à vie. Pendant son incarcération, il subit une évaluation psychologique comprenant le MMPI-2 (voir HYMAN, CALDWELL & NICHOLS, 2013). La comparaison entre le profil des échelles cliniques du MMPI-2 et le profil des échelles RC est remarquable d'enseignements sur la validité de ces deux types de profil.

1. Le profil des échelles cliniques psychopathologiques du MMPI-2 est caractéristique d'un pattern d'inadaptation psychologique de longue durée. Ce sont les échelles **4(Pd)** et **6(Pa)** qui dominent le profil, juste au-dessus de la limite de tolérance de 65. L'échelle **2(D)** se situe à une cote T de 64, donc tout juste en dessous de la limite de tolérance. Les échelles **1(Hs)**, **7(Pt)** et **8(Sc)** avoi-

sinent la cote T de 60. L'échelle **3(Hy)** se situe à une cote T de 50, et l'échelle **9(Ma)** à une cote T tout près de 30. Ce code-type **4-6** suggère un trouble de personnalité paranoïde. Il correspond aux résultats de recherches antérieures chez les meurtriers en série (CRAIG, 2008; NICHOLS, 2006a).

2. Par comparaison, le profil aux échelles RC présente une seule pointe significative, soit $T = 75$ à $RC2$ (*Low Positive Emotions*), suggérant une anhédonie dysphorique d'intensité faible à modérée. L'échelle $RC1$ avoisine une cote T de 60. Les sept autres échelles RC (y compris $RC4$ et $RC6$) sont inférieures à une cote T de 50 (sauf pour $RC3$, qui lui est égale). Les échelles RC n'ont donc pas décelé d'état de santé mentale sévère ni des problèmes de comportement qui apparaissent au profil des échelles cliniques du MMPI-2 (code-type **4-6**) et qui se sont manifestés dans la vie du sujet par un désordre important de pensée paranoïde ainsi que par un fort potentiel de passage à l'acte et un système de croyance tordu. Ainsi, ce manque de sensibilité des échelles RC à la psychopathologie s'ajoute aux autres indications de la littérature clinique, ce qui constitue un handicap réel dans le domaine clinique, voire critique dans le domaine psycholégal.

BUTCHER, HASS, GREENE et NELSON (2015) soulignent deux sources d'insensibilité des échelles RC . D'abord, les huit échelles restructurées RC contiennent beaucoup moins d'items que les huit échelles cliniques psychopathologiques du MMPI-2 (168 versus 411 items), une perte d'information qui entraîne avec elle une perte substantielle de fidélité.¹⁰ Ensuite, ces échelles cliniques sont des mesures hétérogènes comprenant plusieurs types différents de contenu d'items. Ainsi, aux échelles cliniques du MMPI-2, les items ont été retenus parce qu'ils prédisaient le comportement externe, et non parce qu'ils étaient congruents avec d'autres items d'une même échelle. Une faible cohérence interne n'empêche pas nécessairement une bonne prédiction des problèmes de santé mentale. Autrement dit, les items d'une échelle en particulier n'ont pas à se conformer à un seul facteur pour être utile dans la prédiction. En comparaison, la construction des échelles RC a été dictée par une stratégie qui se concentrait sur des dimensions unitaires ou simples plutôt qu'être guidée par un souci de prédiction comportementale.

Cela évoque la «théorie des facettes» de Louis Guttman (GUTTMAN, 1954; GUTTMAN & GREENBAUM, 1998). Une facette est un élément de contenu simple, tandis qu'un concept ou construit psychométrique est un ensemble regroupant deux ou plusieurs facettes, celles-ci pouvant être statistiquement indépendantes l'une de l'autre (d'après

10. Le quotient de perte d'items étant d'environ 0,41, une échelle à fidélité r_{xx} de 0,80 verrait sa fidélité (et précision de mesure) abaissée à environ 0,62.



Guttman). Une intra-corrélation d'échelle très élevée (ou «consistance interne») indique que les facettes incluses sont très présentes, et seulement elles. La question qui se pose est : le comportement de quelqu'un n'est-il déterminé que par un seul concept simple, ne dépend-il que d'un seul facteur ?

L'approche empirique (par association d'items sur des groupes typiques) capte plutôt la configuration de facettes, indépendantes ou non, qui représente le mieux le type psychopathologique visé et est aussi le plus susceptible d'enclencher ou de faire apparaître la psychopathologie. Les «facettes» ou items inclus ne sont pas forcément corrélés ni évidemment compatibles. Cet apport de Guttman constitue certainement un argument de poids en faveur de l'approche empirique, basée sur la réalité clinique observée, plutôt que de la stratégie factorielle. Pour FRIEDMAN & coll. (2015), les causes de ce manque de sensibilité des échelles RC ne sont pas difficiles à trouver. De la même façon qu'aux échelles de contenu du MMPI-2, la plupart des items des échelles RC ont un contenu évident, donc facilement repérable quand on veut présenter de soi-même un portrait factice et avantageux. Qui plus est, les réponses-clés (ou significatives de problèmes) sont du côté VRAI pour la majorité des items, soit 297/366 (81%) des items des 15 échelles de contenu du MMPI-2 et 126/168 (75%) des 8 échelles RC. Ce pattern déséquilibré rend plus facile la «minimisation» de la psychopathologie pour le sujet, en évitant de répondre VRAI. Mais cette stratégie fallacieuse s'avère moins efficace aux échelles cliniques psychopathologiques du MMPI-2, puisque la réponse-clé est VRAI pour 228/411 items soit 55%, une disposition beaucoup plus équilibrée.

Critique 4 : importance factice des concepts «contemporains» de démoralisation et d'anhédonie. Les auteurs des échelles RC (TELLEGEN & coll., 2003) ont poursuivi un deuxième objectif de «croissance» (growth change), secondaire celui-là puisque énoncé en 2006 en réponse à Rogers et coll., 2006 (voir TELLEGEN & coll., 2006). Cet objectif secondaire vise à relier directement le MMPI-2 aux conceptions contemporaines de la psychopathologie. Les auteurs des échelles RC cherchent donc à formuler leurs efforts comme des tentatives de modernisation du MMPI-2 en le mettant en accord avec les conceptions (modèles, connaissances) de la personnalité et de la psychopathologie. Leur raisonnement est le suivant : si les lignes directrices interprétatives se concentrent presque exclusivement sur des pistes de corrélats empiriques et ne fournissent pas de principes conceptuels d'organisation, le MMPI-2 sera laissé en dehors du courant dominant de la réflexion actuelle sur la personnalité et sur la psychopathologie. Le MMPI-2 sera de plus en plus déconnecté des développements conceptuels dans le domaine. Devant cet échec prophétisé du MMPI-2, les échelles RC apparaissent donc comme un

remède.

À cela, RANSON & coll. (2009) répondent que ces soi-disant conceptions contemporaines de la personnalité et de la psychopathologie n'existent pas en dehors du modèle d'affect de Watson et Tellegen. Et pour que les échelles RC servent de rempart à l'obsolescence présumée du MMPI-2, il faudrait que ce modèle soit largement accepté, ce qui n'est pas le cas. En somme, les échelles RC ne sont qu'un autre ensemble d'échelles unidimensionnelles de contenu évident sans pertinence contemporaine particulière. Les deux concepts soi-disant nouveaux prônés par les auteurs des échelles RC, soit la démoralisation et l'anhédonie, n'ont donc ni la portée, ni l'importance qu'ils suggèrent quant au déclassement prophétisé du MMPI-2.

Démoralisation

Les auteurs des échelles RC ont spécifié que la démoralisation est essentiellement une étiquette de recharge pour le pôle PU (*Pleasantness* versus *Unpleasantness*) qui se situe au sommet du modèle hiérarchique de l'affect de Watson et Tellegen. Ils ont souligné sa spécificité théorique et conceptuelle. Pourtant, son application au domaine de la personnalité et de la psychopathologie générale est incertaine et contestée (RANSON & coll., 2009). TELLEGEN & coll. (2006) tentent de faire le pont avec certains auteurs (FRANK & FRANK, 1991; JOINER, WALKER, PETTIT, PEREZ & CUKROWICZ, 2005), qui ont utilisé le même mot *demoralization*. Mais on ne retrouve chez ces auteurs aucun indice qu'ils soient même informés de l'existence du modèle hiérarchique d'affect de Watson et Tellegen. Chez Frank, par exemple, *demoralization* correspond à un concept psychopathologique très large servant à indiquer l'état psychologique partagé dans toutes les psychothérapies et qui amène le patient à chercher de l'aide psychothérapique, peu importe son état émotionnel spécifique. La tentative de faire ce pont avec des auteurs qui ont utilisé le même mot *demoralization* semble donc factice.

Anhédonie

Les auteurs des échelles RC considèrent l'anhédonie comme une construction interchangeable avec le pôle d'activation positive (inversé) de la dimension «activation de l'affect» du modèle de Watson et Tellegen, et équivalente à leur échelle de «Faibles émotions positives RC2» (voir plus haut, Origines des échelles RC). Cela ferait partie des préoccupations de croissance des auteurs d'avoir enfin une mesure d'anhédonie (RC2) pour faire prendre au MMPI-2 le pas des développements contemporains de la personnalité et de la psychopathologie. On semble ici ignorer que dès 1970, C. G. WATSON, KLETT et LOREI (1970) ont élaboré pour le MMPI une mesure



de l'anhédonie (échelle *Anhed*), qui a précédé les travaux de CHAPMAN, CHAPMAN et RAULIN (1976) sur la mesure de l'anhédonie physique et de l'anhédonie sociale, et de KWAPIL, CHAPMAN et CHAPMAN (1999), KWAPIL, BARRANTES-VIDAL et SILVIA (2008) sur l'instrument de mesure *Wisconsin Schizotypy Scales*. Puis, toujours au MMPI, en 1984, Dworkin et Saczynski ont élaboré une mesure de capacité hédonique (échelle *HedCap*). Actuellement, au MMPI-2, l'échelle **INTR** (*Introversión*, la cinquième des échelles **PSY-5**) est elle-même une mesure d'anhédonie (en particulier les sous-échelles **INTR1**, soit «désengagement-anhédonie», et **INTR3**, soit «faible diligence-hypomanie»). De plus, l'échelle **0(Si)** mesure quelque chose d'approchant.¹¹ Dans RANSON & coll. (2009), le tableau 7.4 rend compte des résultats de deux recherches sur les corrélations entre les échelles d'anhédonie physique et sociale de Chapman et les mesures d'anhédonie aux MMPI/MMPI-2 (**RC2**, **INTR** et **Si**). Ces résultats indiquent que **RC2** ne prédit pas mieux, et dans la plupart des cas prédit moins bien que les échelles **0(Si)** et **INTR**. Il est donc faux que les MMPI/MMPI-2 aient pris du retard sur les développements contemporains de la personnalité et de la psychopathologie, du moins en ce qui concerne l'anhédonie.

Récapitulation sur les échelles RC

En somme, avec les échelles **RC**, le fait de rendre celles-ci statistiquement et conceptuellement plus cohérentes et distinctes les unes des autres n'a pas renforcé leur validité externe, ni produit de mesures utiles pour de nombreuses psychopathologies rencontrées dans la pratique clinique (voir GORDON, 2006). Ceci aurait pu être le cas pour des groupes unitaires de symptômes comme l'anxiété ou la colère, mais non pas pour des conditions cliniques complexes comme celles que les huit échelles cliniques psychopathologiques du MMPI-2 tentent de cerner.

Les auteurs des échelles **RC** (TELLEGEN & coll., 2003), comme nous l'écrivions plus haut, ont annoncé : «Les échelles **RC** ont été conçues pour préserver les propriétés descriptives importantes des échelles cliniques existantes du MMPI-2 en améliorant leur caractère distinctif». Cet énoncé rend bien compte des deux objectifs principaux du projet **RC**, soit répondre à un besoin de changement dû à une «carence» alléguée, et répondre à un besoin de «développement». Le premier point concerne la réduction de la covariation compromettant la validité discriminante des échelles cliniques du MMPI-2, ce qui représentait une lacune déjà reconnue et qui n'avait pas besoin de justification. Quant au deuxième point, les auteurs ont cherché

à identifier pour chacune des huit échelles cliniques du MMPI-2 son noyau définitionnel distinct afin de mieux représenter le construit conceptuellement significatif et cliniquement important de l'échelle. Or, la très grande majorité des experts américains des MMPI/MMPI-2 sont d'avis que le projet **RC** a clairement raté ses deux objectifs. L'une des raisons mentionnées est l'adoption par les auteurs d'un modèle d'affect (le modèle Watson-Tellegen) élaboré par l'auteur principal des échelles **RC**, adoption qui a probablement compromis la réussite du projet.

Solutions de rechange au problème de la covariation des échelles cliniques du MMPI-2

Les problèmes que représente la covariation importante entre les échelles cliniques des MMPI/MMPI-2, ainsi que les approches simples offertes pour y remédier, sont reconnus depuis au moins 50 ans. Le plus expéditif serait d'éliminer les items qui chevauchent deux ou plusieurs de ces échelles cliniques. Une telle stratégie répond directement aux objectifs annoncés par les auteurs des échelles **RC**, qui n'ont cependant pas mentionné d'anciennes tentatives en ce sens (voir WELSH, 1952; ADAMS & HORN, 1965).

Tout en reconnaissant le problème de chevauchement des items, TELLEGEN & coll. (2003) semblent finalement le minimiser, mettant en cause d'autres facteurs, telles la variance de style de réponse (p. ex. désirabilité sociale), la comorbidité imprévue (p. ex. certaines fortes corrélations entre le névrotisme et le psychotisme), ou encore l'invalidité (supposée) des «items subtils» (bien que ces items, à la différence des «items évidents», présentent un taux de chevauchement beaucoup plus faible).¹² Cependant, rien n'a été fait pour explorer la force relative des facteurs concurrents, ni pour évaluer l'étendue du rôle joué par le chevauchement lui-même d'items.

La simple stratégie pour pallier la covariation extensive de ces échelles cliniques en éliminant ou en réduisant le chevauchement des items a donc finalement été ignoré par les auteurs des échelles **RC**, malgré qu'elle représentait pour eux un moyen d'atteindre leurs objectifs.

À cet égard, NICHOLS (2006b) a exploré trois solutions différentes au problème, ce en utilisant un échantillon clinique de 26 118 hommes et de 26 425 femmes (CALDWELL, 1997) et leurs résultats au MMPI-2.

1. Dans un premier temps, Nichols a simplement éliminé de chacune des huit échelles cliniques psychopathologiques du MMPI-2 les 35 items qui chevauchent au moins trois de ces échelles de base. Comme conséquence, l'intercorrélation moyenne entre ces

11. L'échelle **0** ou **Si** (*Social Introversion*) est une échelle clinique non-psychopathologique faisant partie du profil clinique de base des MMPI/MMPI-2.

12. La prétendue invalidité des items «subtils» constitue une affirmation tout à fait gratuite, car il s'agit d'items qui ont été retenus au MMPI, ayant franchi avec succès la stratégie des «groupes contrastés», d'abord une première fois, puis par validation croisée (voir au début du présent texte, section *Rappel historique : stratégie psychométrique empirique*). Les auteurs des échelles **RC** ne peuvent évidemment pas accepter ce fait incompatible avec leur stratégie théorique-factorielle.



échelles est passée de 0,59 à 0,39 (diminution de 20% de la variance partagée). Pourtant, la corrélation moyenne entre l'échelle modifiée et l'échelle-mère intacte est restée élevée (à 0,94), une perte de seulement 13% de la variance partagée.

2. Dans un deuxième temps, Nichols a identifié 37 items des huit échelles cliniques dont la contribution est apparue significative par rapport au facteur 1 dans deux analyses indépendantes (JOHNSON & coll., 1984; WALLER, 1999). Il a ensuite supprimé la variance de cet ensemble de 37 items de chacune des huit échelles cliniques. En utilisant cette stratégie, Nichols a obtenu une augmentation d'indépendance d'échelle de 12% (diminution des intercorrélations moyennes de 0,59 à 0,47), avec maintien d'une corrélation moyenne de 0,97 des échelles modifiées avec leur échelle-mère, et une perte de seulement 6% de la variance partagée.
3. Dans un troisième temps, Nichols a combiné les deux premières stratégies en supprimant des échelles cliniques du MMPI-2 les items qui apparaissaient sur les deux options de «marqueur». Il a obtenu une augmentation d'indépendance d'échelle de 25% (diminution des intercorrélations moyennes de 0,59 à 0,32), tout en maintenant une corrélation moyenne de 0,88, soit une perte de 23% de la variance partagée.

Ces tests simples montrent tous comment des ajustements peu compliqués aux échelles cliniques du MMPI-2 peuvent augmenter leur indépendance relative sans perte importante pour la conformité des échelles modifiées avec leurs échelles-mères. Ils démontrent avec quelle facilité les objectifs exprimés par les auteurs des échelles RC pouvaient être atteints, sans recours à des procédures radicales qui ont abouti à un ensemble d'échelles RC dotées de concordances à leurs échelles-mères trop faibles pour justifier une interprétation significative de ces nouvelles échelles (voir ROGERS, SEWELL, HARRISON & JORDAN, 2006; BUTCHER & coll., 2006).

Nouvelle analyse de la covariation des échelles cliniques du MMPI-2

Les auteurs des échelles RC ont rejeté les analyses de Nichols (voir ci-dessus : *Solutions de rechange...*) pour deux motifs (voir TELLEGEN & coll., 2006). Ils ont jugé que des analyses uniquement internes des MMPI/MMPI-2 étaient insuffisantes sans d'autres analyses impliquant des critères externes, et ils ont également jugé que les conclusions de NICHOLS (2006b) et de ROGERS & coll. (2006) étaient basées sur un ensemble inadéquat de données.

À ces critiques, on répond d'abord (RANSON & coll.,

2009) que la valeur des analyses externes ne diminue en rien l'importance des conclusions résultant des analyses internes et qu'en fait, la méthode et les procédures suivies dans la construction des échelles RC elles-mêmes dépendaient en grande partie de telles analyses internes.

Ensuite, les critiques ont porté sur le caractère inadéquat de l'échantillon de CALDWELL (1997), échantillon supposé ne représenter aucun milieu clinique non plus qu'une autre population correctement définie, mais plutôt un échantillon composite recruté au hasard et mal défini.

À cela, Ranson, Nichols, Rouse et Harrington répondent (2009) que les auteurs des échelles RC, dont Tellegen et Ben-Porath, malgré les disponibilités qui leur étaient offertes, ont opté de ne refaire aucune des analyses de Nichols ni d'utiliser plutôt leurs propres échantillons auxquels ils sont seuls à avoir accès. De telles analyses répliquées, si elles avaient été présentées, auraient pu fournir une preuve directe des insuffisances présumées de l'échantillon Caldwell. Compte tenu de l'ensemble des données cliniques de Caldwell ($N > 52000$), il semble toutefois peu probable que les valeurs rapportées à partir des analyses de Nichols soient dépourvues de valeur.

Ainsi, pour répondre aux préoccupations pourtant faciles à contrer des auteurs des échelles RC, RANSON & coll. (2009) ont voulu apporter de nouvelles analyses basées sur 24 de leurs 25 échantillons diversifiés, en excluant donc les données de Caldwell. Il s'agissait d'échantillons obtenus (ROUSE & coll., 2008) dans des milieux cliniques, médico-légaux, universitaires, médicaux, d'emploi, militaires et communautaires ($N = 29983$). Les auteurs (ibid., tableau 7.2) rapportent que les résultats de cette nouvelle analyse reproduisent largement les conclusions antérieures de NICHOLS (2006b).

Le petit gain d'indépendance (c.-à-d. la diminution de la covariation) obtenu par TELLEGEN & coll. (2003) pour les échelles RC est anéanti par une perte de la variance partagée avec les échelles cliniques du MMPI-2. En effet, les corrélations entre les échelles RC dans les échantillons utilisés par les auteurs des échelles RC contestataires ne montrent que des réductions minimales de la covariation (5%), mais au prix d'une perte de conformité d'environ 48% avec leurs échelles-mères.¹³ Donc, l'objectif premier des auteurs des échelles RC, c'est-à-dire de réduire considérablement la covariation qui compromet la validité discriminante des échelles cliniques MMPI-2, semble n'avoir été atteint que marginalement, et par des méthodes à la fois plus discutables et compliquées alors que des moyens plus simples suffisaient.

De tous ces auteurs qui ont expérimenté de nouvelles

13. Le terme anglais *fidelity* utilisé par RANSON & coll. (2009), inconnu en jargon psychométrique anglo-saxon, est ambigu, ne correspondant pas au terme français «fidélité» ou *reliability* en anglais. Le contexte de la phrase nous fait opter pour le terme français «conformité», c.-à-d. (cor)relation ou proximité de deux versions du même test ou deux éditions d'une même échelle.



façons de diminuer la covariation de ces échelles cliniques, aucun n'a eu l'autorisation d'intervenir constructivement sur le test MMPI-2, sauf Tellegen, Ben-Porath et coll., lesquels sont en fait les «protégés» de la maison d'édition *The University of Minnesota Press* qui contrôle ce test.

Conclusion : passage raté des échelles RC au MMPI-2-RF

En 2003, avec l'assentiment de la maison d'édition *The University of Minnesota Press*, les échelles RC ont été introduites dans le *MMPI-2 Extended Score Report* par le service de correction informatisée *NCS-Pearson Assessments*. On peut concevoir que cette insertion fut comme un chien dans un jeu de quilles, étant donné le peu de parenté psychométrique qu'ont ces échelles avec les instruments MMPI et MMPI-2. Pourtant, ce sont plusieurs auteurs du MMPI-2 (Tellegen, Ben-Porath, Graham : voir BUTCHER & coll., 1989; BUTCHER & coll., 2001) qui ont été à l'origine de cette insertion (voir TELLEGEN & coll., 2003). Au contraire, Butcher n'a jamais lâché prise quant à ses critiques des échelles RC, soit comme directeur de publication (*Oxford Handbook of Personality Assessment* : voir ROUSE & coll., 2008), soit comme auteur (voir par exemple BUTCHER & coll., 2006; BUTCHER & coll., 2015).

Puis, en 2008 (voir BEN-PORATH & TELLEGEN, 2008, 2011; BEN-PORATH, 2013), fut édité le MMPI-2-RF (RF pour *Restructured Form*), qui consiste en une structure factorielle hiérarchique (*top-down approach*), incluant au niveau intermédiaire les échelles RC, un inventaire auto-rapporté de 338 items regroupés en 51 échelles nouvelles ou révisées. Il ne s'agit pas d'une révision du MMPI-2 (voir BUTCHER & coll., 2015; FRIEDMAN & coll., 2015; GRAHAM, 2012; GREENE, 2011; NICHOLS, 2011; RUCHENNE, 2019); il en est plutôt un avatar et il diffère essentiellement du MMPI-2. Cette forme RF doit donc être considérée comme un nouvel instrument plutôt qu'une simple mise à jour du MMPI-2, tel qu'avait été le passage du MMPI au MMPI-2. En somme, le MMPI-2-RF est un instrument psychométrique nouveau et non validé sur plusieurs plans, qui n'a pas le support empirique (avec la crédibilité qui s'y rattache), ni les données interprétatives du MMPI-2 (FRIEDMAN & coll., 2015). L'appellation MMPI-2(-RF) est même considérée comme indue, voire illégitime.

Pour ajouter l'insulte à l'affront, les éditeurs du test *University of Minnesota Press* et leurs distributeurs *Pearson Assessments* ont récemment mis sur le marché un «MMPI-3» (produit issu du MMPI-2-RF), qui sur le fond n'a non plus rien à voir avec les MMPI/MMPI-2.

Pour l'instant, les échelles RC, issues d'une stratégie théorique et factorielle, ne font pas le poids vis-à-vis de la stratégie psychométrique empirique qui a fait (et fait encore) des instruments MMPI/MMPI-2 le *nec plus ultra* des inventaires de personnalité utilisés dans les contextes cli-

niques et psycho-légaux. Si, en principe, il est intéressant qu'un inventaire de personnalité mette en lumière pour l'utilisateur les construits factoriels qui le constituent, il vise d'abord et surtout à produire des données probantes et des indications fiables pour la pratique, que celle-ci soit clinique ou psycholégale : c'est ce qui permet l'approche empirique du MMPI ou MMPI-2 classique. L'identification d'un syndrome ou d'un trouble clinique à partir d'un critère externe objectif est ici l'élément essentiel de la stratégie psychométrique empirique. De nos jours, elle procède typiquement par l'analyse en fonctions discriminantes, donnant lieu à une forme de validité qu'on nomme aussi «discriminante» (MCLACHLAN, 2004). L'analyse discriminante permet, par association d'items sur des groupes typiques, de capter la configuration de «facettes» (GUTTMAN, 1954; GUTTMAN & GREENBAUM, 1998) qui constituent chaque type et sont susceptibles d'enclencher ou de faire apparaître la psychopathologie. Ce sont les corrélations des items avec un critère externe qui importent, et non les corrélations entre les items.

Si l'on entreprenait une re-validation des échelles cliniques des MMPI/MMPI-2 en formant de nouveaux groupes cliniques pour comparaison avec un échantillon normal, ce qui n'a jamais été refait depuis les années 1940, peut-être l'analyse en fonctions discriminantes représenterait-elle la solution de l'avenir. En attendant, une disparition éventuelle du MMPI-2 serait le fait d'une psychométrie mal pensée et de manœuvres éditoriales douteuses, avec des répercussions cliniques désastreuses.

Remerciements

Nous tenons à remercier Louis Laurencelle du temps précieux qu'il nous a accordé pour son support et pour la révision du présent texte.

Références

- ADAMS, D. K. & HORN, J. L. (1965). Nonoverlapping keys for the MMPI scales. *Journal of Consulting Psychology*, 29(3), 284-284. doi :10.1037/h0022112
- AMERICAN PSYCHIATRIC ASSOCIATION. (1952). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC : Author.
- ANDRÉ, N., LOYE, N. & LAURENCELLE, L. (2015). La validité psychométrique : un regard global sur le concept centenaire, sa genèse, ses avatars. *Mesure et évaluation en éducation*. 37(3), 125-148. doi :10.7202/1036330ar
- ARCHER, E. M., HAGAN, L. D., MASON, J., HANDEL, R. & ARCHER, R. P. (2012). MMPI-2-RF characteristics of custody evaluation litigants. *Assessment*, 19(1), 14-20. doi :10.1177/1073191110397469
- BEN-PORATH, Y. S. (2013). *Interpreting the MMPI-2-RF*. Minneapolis : University of Minnesota Press.



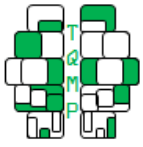
- BEN-PORATH, Y. S. & SHERWOOD, N. E. (1993). *The MMPI-2 content component scales*. Minneapolis : University of Minnesota Press.
- BEN-PORATH, Y. S. & TELLEGEN, A. (2008). *MMPI-2-RF : Manual for administration, scoring and interpretation*. Minneapolis : University of Minnesota Press.
- BEN-PORATH, Y. S. & TELLEGEN, A. (2011). *Minnesota Multiphasic Personality Inventory-2 Restructured Form : Manual for administration, scoring, and interpretation*. Minneapolis, MN : University of Minnesota Press.
- BERG, I. A. (1959). The unimportance of test item content, p. 83-99. In *Objective Approaches to Personality Assessment*. Princetown, NJ : D. Van Nostrand Company inc.
- BOLINSKEY, P. K. & NICHOLS, D. S. (2011). Construct drift in the MMPI-2 Restructured Clinical (RC) scales : further evidence and a possible historic example. *Journal of Clinical Psychology*, 67(9), 907-917. doi :[10.1002/jclp.20814](https://doi.org/10.1002/jclp.20814)
- BUTCHER, J. N. (2011). *A beginner's guide to the MMPI-2, 3rd Ed.* Washington, DC : American Psychological Association Press.
- BUTCHER, J. N., DAHLSTROM, W. G., GRAHAM, J. R., TELLEGEN, A. & KAEMMER, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2) : Manual for administration and scoring*. Minneapolis, MN : University of Minnesota Press.
- BUTCHER, J. N., GRAHAM, J. R. & BEN-PORATH, Y. S. (1995). Methodological problems and issues in MMPI, MMPI-2, and MMPI-A research. *Psychological Assessment*, 7(3), 320-329.
- BUTCHER, J. N., GRAHAM, J. R., BEN-PORATH, Y. S., TELLEGEN, A., DAHLSTROM, W. G. & KAEMMER, B. (2001). *MMPI-2 Manual for administration and scoring. Éd. rév.* Minneapolis : University of Minnesota Press.
- BUTCHER, J. N., GRAHAM, J. R., WILLIAMS, C. L. & BEN-PORATH, Y. S. (1990). *Development and uses of the MMPI-2 content scales*. Minneapolis : University of Minnesota Press.
- BUTCHER, J. N., HAMILTON, C. K., ROUSE, S. V. & CUMELA, E. J. (2006). The deconstruction of the Hy Scale of MMPI-2 : Failure of RC3 in measuring somatic symptom expression. *Journal of Personality Assessment*, 87(2), 186-192. doi :[10.1207/s15327752jpa8702_08](https://doi.org/10.1207/s15327752jpa8702_08)
- BUTCHER, J. N., HASS, G. A., GREENE, R. L. & NELSON, L. D. (2015). *Using the MMPI-2 in Forensic Assessment*. doi :[10.1037/14571-001](https://doi.org/10.1037/14571-001)
- BUTCHER, J. N. & TELLEGEN, A. (1978). *Essentials of the MMPI-2 and MMPI-A clinical interpretation (2nd ed.)* Minneapolis, MN : University of Minnesota Press.
- BUTCHER, J. N., WILLIAMS, C. L., GRAHAM, J. R., TELLEGEN, A., BEN-PORATH, Y. S., ARCHER, R. P. & KAEMMER, B. (1992). *Manual for administration scoring and interpretation of the Minnesota Multiphasic Personality Inventory for Adolescents : MMPI-A*. Minneapolis : University of Minnesota Press.
- CALDWELL, A. B. (1997). *MMPI-2 data research file for clinical patients*.
- CALDWELL, A. B. (2006). Maximal measurement or meaningful measurement : the interpretive challenges of the MMPI-2 Restructured Clinical (RC) scales. *Journal of Personality Assessment*, 87(2), 193-201. doi :[10.1207/s15327752jpa8702_09](https://doi.org/10.1207/s15327752jpa8702_09)
- CARROLL, J. M., YIK, M. S. M., RUSSELL, J. A. & FELDMAN-BARRETT, L. (1999). On the psychometric principles of affect. *Review of General Psychology*, 3(1), 14-22.
- CHAPMAN, L. J., CHAPMAN, J. P. & RAULIN, M. L. (1976). Scales for physical and social anhedonia. *Journal of Abnormal Psychology*, 85(4), 374-382. doi :[10.1037/0021-843X.85.4.374](https://doi.org/10.1037/0021-843X.85.4.374)
- CRAIG, R. J. (2008). MMPI based forensic-psychological assessment of lethal violence. In H. V. HALL (Éd.), *Forensic psychology and neuropsychology for criminal and civil cases (p (p. 393-416)*. Boca Raton, FL : CRC Press.
- DWORKIN, R. H. & SACZYNSKI, K. (1984). Individual differences in hedonic capacity. *Journal of Personality Assessment*, 48(6), 620-626. doi :[10.1207/s15327752jpa4806_8](https://doi.org/10.1207/s15327752jpa4806_8)
- FRANK, J. D. & FRANK, J. B. (1991). *Persuasion and healing : a comparative study of psychotherapy (3rd ed.)* Baltimore, MD : Johns Hopkins University Press.
- FRIEDMAN, A. F., BOLINSKEY, P. K., LEVAK, R. W. & NICHOLS, D. S. (2015). *Psychological Assessment with the MMPI-2/MMPI-2-RF. Third Edition*. doi :[10.4324/9780203119556](https://doi.org/10.4324/9780203119556)
- FRIEDMAN, A. F., GLESER, G. C., SMELTZER, D. J., WAKEFIELD, J. A., Jr. & SCHWARTZ, M. S. (1983). MMPI overlap item scales for differentiating psychotics, neurotics, and nonpsychiatric groups. *Journal of Consulting and Clinical Psychology*, 51(4), 629-631. doi :[10.1037/0022-006X.51.4.629](https://doi.org/10.1037/0022-006X.51.4.629)
- GORDON, R. M. (2006). False assumptions about psychopathology, hysteria and the MMPI-2 Restructured Clinical Scales. *Psychological Reports*, 98(3), 870-872.
- GORDON, R. M., STOFFEY, R. W. & PERKINS, B. L. (2013). Comparing the sensitivity of the MMPI-2 clinical scales and the MMPI-RC scales to clients rated as psychotic, borderline or neurotic on the Psychodiagnostic Chart. *Psychology*, 4(9B), 12-16.
- GRAHAM, J. R. (2012). *MMPI-2 : Assessing psychopathology and personality (4th ed.)* New York, NY : Oxford University Press.
- GRAHAM, J. R., BEN-PORATH, Y. S. & McNULTY, J. L. (1999). *MMPI-2 correlates for outpatient community mental*



- health settings. Minneapolis : University of Minnesota Press.
- GREEN, D. P., SALOVEY, P. & TRUAX, K. M. (1999). Static, dynamic, and causative bipolarity of affect. *Journal of Personality and Social Psychology*, 76(5), 856-867.
- GREENE, R. L. (1991). *The MMPI-2/MMPI : An interpretive manual*. Boston : Allyn & Bacon.
- GREENE, R. L. (2000). *The MMPI-2 : An interpretive manual — 2nd Edition*. Boston, MA : Allyn & Bacon.
- GREENE, R. L. (2011). *The MMPI-2/MMPI-2-RF : An interpretive manual — 3rd Edition*. Boston, MA : Allyn & Bacon.
- GREENE, R. L., ROUSE, S. V., BUTCHER, J. N., NICHOLS, D. S. & WILLIAMS, C. L. (2009). The MMPI-2 Restructured Clinical (RC) Scales and redundancy : response to Tellegen, Ben-Porath, and Sellbom. *Journal of Personality Assessment*, 91(3), 222-226. doi :[10 . 1080 / 00223890902800825](https://doi.org/10.1080/00223890902800825)
- GUTTMAN, L. (1954). An outline of some new methodology for social research. *The Public Opinion Quarterly*, 18, 395-404.
- GUTTMAN, L. & GREENBAUM, C. W. (1998). Facet theory : its development and current status. *European Psychologist*, 3(1), 3-36.
- HARKNESS, A. R., McNULTY, J. L. & BEN-PORATH, Y. S. (1995). The Personality Psychopathology Five (PSY-5) : Constructs and MMPI-2 scales. *Psychological Assessment*, 7, 104-114. doi :[10.1037/1040-3590.7.1.104](https://doi.org/10.1037/1040-3590.7.1.104)
- HATHAWAY, S. R. & MCKINLEY, J. C. (1940). A multiphasic personality schedule (Minnesota) : I. Construction of the schedule. *Journal of Psychology*, 10, 249-254.
- HATHAWAY, S. R. & MCKINLEY, J. C. (1942). A multiphasic personality schedule (Minnesota) : III. The measurement of symptomatic depression. *Journal of Psychology*, 14, 73-84.
- HATHAWAY, S. R. & MCKINLEY, J. C. (1943). *The Minnesota Multiphasic Personality Inventory Manual*. Minneapolis, MN : University of Minnesota Press.
- HELMES, E. & REDDON, J. R. (1993). A perspective on developments in assessing psychopathology : A critical review of the MMPI and MMPI-2. *Psychological Bulletin*, 113(3), 453-471. doi :[10.1037/0033-2909.113.3.453](https://doi.org/10.1037/0033-2909.113.3.453)
- HYMAN, E. J., CALDWELL, A. & NICHOLS, D. S. (2013). The iterations of the MMPI-reliability, validity, and homicide, San Diego, CA : Meeting of the Society for Personality Assessment.
- JOHNSON, J. H., NULL, C., BUTCHER, J. N. & JOHNSON, K. N. (1984). Replicated item level factor analysis of the full MMPI. *Journal of Personality and Social Psychology*, 47(1), 105-114. doi :[10.1037/0022-3514.47.1.105](https://doi.org/10.1037/0022-3514.47.1.105)
- JOINER, T. E., Jr., WALKER, R. L., PETTIT, J. W., PEREZ, M. & CUKROWICZ, K. C. (2005). Evidence-based assessment of depression in adults. *Psychological Assessment*, 17(3), 267-277.
- KWAPIL, T. R., BARRANTES-VIDAL, N. & SILVIA, P. J. (2008). The dimensional structure of the Wisconsin Schizotypy Scales : factor identification and construct validity. *Schizophrenia Bulletin*, 34(3), 444-457.
- KWAPIL, T. R., CHAPMAN, L. J. & CHAPMAN, J. (1999). Validity and usefulness of the Wisconsin manual for assessment psychotic-like experiences. *Schizophrenia Bulletin*, 25(2), 363-375.
- LARSEN, R. J. & DIENER, E. (1992). Promises and problems with the circumplex model of emotion. In M. S. CLARK (Éd.), *Review of personality and social psychology* (p. 25-59). Inc : No. 13. Emotion (p.. Sage Publications.
- LAURENCELLE, L. (1998). *Théorie et techniques de la mesure instrumentale*. Québec : Presses de l'Université du Québec.
- LEARY, T. (1957). *Interpersonal diagnosis of personality*. New York : Ronald Press.
- MCLACHLAN, G. J. (2004). *Discriminant analysis and statistical pattern recognition*. Interscience : Wiley.
- MEEHL, P. E. (1945a). An investigation of a general normality or control factor in personality testing. *Psychological Monographs*, 4(274), 1-199.
- MEEHL, P. E. (1945b). The dynamics of « structured » personality tests. *Journal of Clinical Psychology*, 1, 296-303.
- MEGARGEE, E. I. (2006). *Using the MMPI-2 in criminal justice and correctional settings*. Minneapolis, MN : University of Minnesota Press.
- NICHOLS, D. S. (2006a). Tell me a story : MMPI responses and personal biography in the case of a serial killer. *Journal of Personality Assessment*, 86(3), 242-262.
- NICHOLS, D. S. (2006b). The trials of separating bath water from the baby : A review and critique of the MMPI-2 Restructured Clinical scales. *Journal of Personality Assessment*, 87(2), 121-138. doi :[10 . 1207 / s15327752jpa8702_02](https://doi.org/10.1207/s15327752jpa8702_02)
- NICHOLS, D. S. (2011). *Essentials of MMPI-2 assessment (2nd ed.)* Hoboken, N.J. : John Wiley & Sons, Inc.
- PARISIEN, M. (1989). Le MMPI-2 : surtout de bonnes nouvelles. *Psychologie Québec*, 6(4), 6-7.
- PARISIEN, M. (1999). Le MMPI et le MCMI en expertise psycholégale. In *L'expertise psycholégale. Balises méthodologiques et déontologiques* (p. 263-308). Québec : Presses de l'Université du Québec.
- PARISIEN, M. (2014). Le MMPI et le MCMI en expertise psycholégale. In *L'expertise psycholégale. Balises méthodologiques et déontologiques, 2ième édition* (p. 273-326). Québec : Presses de l'Université du Québec.



- PIZITZ, T. & McCULLAUGH, J. (2011). An overview of male stalkers' personality profiles using the MMPI-2. *American Journal of Forensic Psychiatry*, 32(3), 31-46.
- RANSON, M. B., NICHOLS, D. S., ROUSE, S. V. & HARRINGTON, J. I. (2009). Changing or replacing an established psychological assessment standard : Issues, goals, and problems with special reference to recent developments in the MMPI-2. In J. N. BUTCHER (Éd.), *Oxford Handbook of Personality Assessment* (p. 112-139). New York : Oxford University Press.
- ROGERS, R., SEWELL, K. W., HARRISON, K. S. & JORDAN, M. J. (2006). The MMPI-2 restructured clinical scales : a paradigmatic shift in scale development. *Journal of Personality Assessment*, 87(2), 139-147. doi :[10.1207/s15327752jpa8702_03](https://doi.org/10.1207/s15327752jpa8702_03)
- ROUSE, S. V., GREENE, R. L., BUTCHER, J. N., NICHOLS, D. S. & WILLIAMS, C. L. (2008). What do the MMPI-2 Restructured Clinical Scales reliably measure? Answers from multiple research settings. *Journal of Personality Assessment*, 90(5), 435-442. doi :[10.1207/s15327752jpa8702_03](https://doi.org/10.1207/s15327752jpa8702_03)
- RUCHENNE, D. (2019). *L'examen clinique de la personnalité avec le MMPI-2 et le MMPI-A*. Fondement et méthode. Préfacé par J. N. Butcher. Bruxelles : Mardaga.
- RUSSELL, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178. doi :[10.1037/h0077714](https://doi.org/10.1037/h0077714)
- SABORIO, C. & HASS, G. (2012). *The Psychological functioning of sexually assaulted women as seen through the MMPI-2 clinical and restructured clinical scales*. Chicago, IL : the meeting of the Society for Personality Assessment.
- SELLBOM, M. & BEN-PORATH, Y. S. (2006). *The Minnesota Multiphasic Personality Inventory-2, p. 19-55*. in ARCHER, R. P. (Dir.) *Forensic Uses of Clinical Assessment Instruments*. New York : Psychology Press.
- STRONG, E. K. (1927). Vocational interest test. *Educational Record*, 8, 107-121.
- TELLEGEN, A. & BEN-PORATH, Y. S. (2011). *Minnesota Multiphasic Personality Inventory-2-Restructured Form : Technical manual*. Minneapolis, MN : University of Minnesota Press.
- TELLEGEN, A., BEN-PORATH, Y. S., McNULTY, J., ARBISI, P., GRAHAM, J. R. & KAEMMER, B. (2003). *MMPI-2 : Restructured Clinical (RC) Scales : Development, validation, and interpretation*. Minneapolis, MN : University of Minnesota Press.
- TELLEGEN, A., BEN-PORATH, Y. S., SELDOM, M., ARBISI, P. A., McNULTY, J. L. & GRAHAM, J. R. (2006). Further evidence on the validity of the MMPI-2 Restructured Clinical (RC) scales : addressing questions raised by Rogers, Sewell, Harrison and Jordan and Nichols. *Journal of Personality Assessment*, 87(2), 148-171. doi :[10.1207/s15327752jpa8702_04](https://doi.org/10.1207/s15327752jpa8702_04)
- TELLEGEN, A., BEN-PORATH, Y. S. & SELLBOM, M. (2009). Construct validity of the MMPI-2 Restructured Clinical (RC) Scales : reply to Rouse, Greene, Butcher, Nichols, and Williams. *Journal of Personality Assessment*, 91(3), 211-221. doi :[10.1080/00223890902794192](https://doi.org/10.1080/00223890902794192)
- TELLEGEN, A., WATSON, D. & CLARK, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, 10(4), 297-303.
- THAYER, R. E. (1996). *The origin of everyday moods*. tension, Managing energy.
- THOMAS, M. L. & YOUNGJOHN, J. R. (2009). Let's not get hysterical : comparing the MMPI-2 validity, clinical, and RC scales in TBI litigants tested for effort. *Clinical Neuropsychologist*, 23(6), 1067-1084. doi :[10.1080/13854040902795000](https://doi.org/10.1080/13854040902795000)
- WALLACE, A. & LILJEQUIST, L. (2005). A comparison of the correlational structures and elevation patterns of the MMPI-2 Restructured Clinical (RC) and Clinical Scales. *Assessment*, 12(3), 290-294. doi :[10.1177/1073191105276250](https://doi.org/10.1177/1073191105276250)
- WALLER, N. G. (1999). Searching for structure in the MMPI. In S. E. EMBRETSON & S. L. HERSHBERGER (Éd.), *The new rules of measurement : What every psychologist and educator should know* (p. 185-217). Mahwah, NJ : Lawrence Erlbaum.
- WATSON, C. G., KLETT, W. G. & LOREI, T. W. (1970). Toward an operational definition of anhedonia. *Psychological Reports*, 26(2), 371-376. doi :[10.2466/pr0.1970.26.2.371](https://doi.org/10.2466/pr0.1970.26.2.371)
- WATSON, D. & TELLEGEN, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98(2), 219-235. doi :[10.1037/0033-2909.98.2.219](https://doi.org/10.1037/0033-2909.98.2.219)
- WATSON, D., WIESE, D., VAIDYA, J. & TELLEGEN, A. (1999). The two general activation systems of affect : structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, 76(5), 820-838. doi :[10.1037/0022-3514.76.5.820](https://doi.org/10.1037/0022-3514.76.5.820)
- WELSH, G. S. (1952). A factor study of the MMPI using scales with item overlap eliminated. *American Psychologist*, 7(1), 341-341.
- WELSH, G. S. (1956). Factor dimensions A and R. In G. S. WELSH & W. G. DAHLSTROM (Éd.), *Basic readings on the MMPI in psychology and medicine* (p. 264-281). Minneapolis, MN : University of Minnesota Press.
- WELSH, G. S. (2000). Factor dimensions A and R. In J. N. BUTCHER (Éd.), *Basic sources on the MMPI-2* (p. 73-92). Minneapolis, MN : University of Minnesota Press.
- WIENER, D. N. (1948). Subtle and obvious keys for the MMPI. *Journal of Consulting Psychology*, 12(3), 164-170. doi :[10.1037/h0055594](https://doi.org/10.1037/h0055594)



Citation

PARISIEN, M. (2021). Validité empirique versus validité factorielle dans les inventaires de personnalité : le MMPI-2 et les échelles restructurées RC.// Empirical validity versus factorial validity in personality inventories : The MMPI-2 and the restructured RC scales. *The Quantitative Methods for Psychology*, 17(2), 87-104. doi :[10.20982/tqmp.17.2.p087](https://doi.org/10.20982/tqmp.17.2.p087)

Copyright © 2021, PARISIEN. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 16/01/2021 ~ Accepted: 03/06/2021