

## EDITORIAL

## US ET ABUS DES ÉCHELLES D'ÉVALUATION EN PSYCHOPATHOLOGIE

## USE AND ABUSE OF RATING SCALES IN PSYCHOPATHOLOGY

D. BOBON \*

## R É S U M É

La quantification de la psychopathologie est devenue indispensable en recherche clinique car, même pour les données de laboratoire, la *signification statistique* n'acquiert de sens que si, par des corrélations avec des données cliniques quantifiées, on lui trouve une *signification clinique*.

Mais il existe déjà trop d'échelles d'évaluation dont la validité n'a pas été suffisamment éprouvée. Il est grand temps que psychiatres et psychologues cliniciens contribuent à des recherches non plus *avec* des échelles mais *sur* ces échelles.

Après un bref historique, l'auteur aborde systématiquement des aspects de la validité, de la fidélité et de la sensibilité des échelles d'évaluation qui lui paraissent importantes et/ou sous-estimées.

**Mots-clés :** Echelles d'évaluation - Psychopathologie quantitative.

## S U M M A R Y

*The quantification of psychopathology with the help of psychiatric rating scales has reached the point of no-return: there is no research in psychiatry without a search for its clinical significance, and no research on this clinical significance without rating scales.*

*But there is a dramatic need for an increase in the number and quality of projects not only **with** rating scales but also **on** rating scales. At present, there exist too many such scales and not enough research on their validity vs. limitations; this research could be the field of a more fruitful collaboration between psychiatrists and clinical psychologists.*

*After a brief survey of the historical background of the quantification of psychopathology (from psychophysics to Wundt, Galton, Kraepelin and the era of psychopharmacology) and of recent literature on the available scales, the author comments on various aspects of quantitative psychopathology which have direct implications on the quality of the measures: self-rating vs. observer ratings, global or specific vs. comprehensive scales, back-translation, use of video in the detection of semantic or conceptual mistranslations, questions and response modalities, incidence of the validation sample on the validity of a scale, orthogonality of the measurements (exclusion of items on anxiety from a depression scale, weighting of item scores), cutoff scores, semi-structured interview, number and criteria of steps of severity, transcultural differences in the pathology threshold, incidence of a time-blind evaluation on the sensitivity of the measurement of change, factor analysis of pre- vs. posttreatment scores etc.*

**Key-words:** *Quantitative psychopathology - Rating scales.*

Reçu le 16 juillet 1987 ; accepté le 1<sup>er</sup> décembre 1987.

\* Chargé de cours de psychiatrie légale et sociale à l'Université de Liège - Service de Neuropsychiatrie (Prof. G. Franck).  
Policlinique neuropsychiatrique CHU de Bavière,  
Boulevard de la Constitution, 66, B-4020 Liège, Belgique.

## De la psychophysique aux "rating scales"

Si c'est le philosophe genevois Charles Bonnet qui, il y a plus de deux siècles, proposa le terme de *psychomètre*, "présumant qu'un jour on mesurera les esprits comme on mesure les corps" (Bonnet 1770), c'est l'anglais Francis Galton qui fut le véritable père de la psychométrie. Dès 1879, il écrivait que "jusqu'à ce que les phénomènes de toute branche du savoir aient été soumis à la mesure et au nombre, cette branche ne peut assumer le statut et la dignité d'une science".

C'est Galton qui a non seulement trouvé les notions statistiques de corrélation et de régression mais qui a également introduit en psychologie la mesure différentielle et le test d'associations d'images ou de mots, précurseur des tests d'association de Kraepelin, repris par Jung, et des tests projectifs de type Rorschach.

La même année que celle où paraissait l'article de Galton sur les "Psychometric experiments", Wundt commençait, dans un local exigü de l'Université de Leipzig, des travaux de psychophysologie qui allaient rapidement dégager les fonctions mentales supérieures de la "psychophysique" des sensations de Weber et de Fechner, montrant qu'il n'y a pas de psychophysique indépendante du psychisme. Les travaux de Wundt ont profondément influencé son compatriote Kraepelin et le psychologue américain James McKeen Cattell qui, après un doctorat à Leipzig et un séjour chez Galton, rentra aux États-Unis pour y publier l'article dans lequel il créa le terme de *test mental* (Cattell 1890).

L'influence de Wundt et de Galton sur Kraepelin et sur James McKeen Cattell - à ne pas confondre avec Raymond B. Cattell, l'auteur du 16 PF et d'une échelle d'anxiété - est rappelée par Pichot dans son rapport au 54<sup>e</sup> congrès de psychiatrie et de neurologie de langue française à Bordeaux (Pichot 1956) et dans son traité sur les "Méthodes psychométriques en clinique" (Delay et al. 1966).

Malgré plusieurs travaux remarquables de rigueur publiés au tournant du siècle, les tentatives de Kraepelin d'intéresser les psychiatres à l'analyse expérimentale et à la codification sinon à la quantification des phénomènes psychopathologiques resteront pratiquement lettre morte jusqu'à l'avènement des psychotropes dans les années cinquante.

En revanche, nous ne nous attarderons pas à rappeler l'engouement initial des psychologues pour la psychométrie, en particulier en France (Binet en tête) et aux USA. C'est ainsi par exemple que la "Visual Analogue Scale" ou VAS, sortie de l'oubli par l'anglais Aitken (Aitken et al. 1965), avait été conçue dès 1920 par Walter D. Scott, Président d'une société d'ingénieurs industriels de Chicago, à qui l'on avait demandé un moyen d'évaluer anonymement, simplement et objectivement les qualités et défauts de candidats à un poste de direction ;

c'est aussi au "Scott Co Laboratory" que semble être née l'expression même de *rating scale*.

Si la première "psychiatric rating scale" contemporaine fut celle du psychologue américain Wittenborn (1950, 1955), l'essor des échelles d'évaluation de la psychopathologie est, on le sait, lié à l'essor des psychotropes. Ce sont donc le plus souvent des psychiatres qui ont conçu les échelles d'évaluation modernes dans une visée de mesure du changement du tableau clinique sous traitement, alors que les tests de personnalité et les "rating scales" antérieures l'avaient été dans le but de mesurer les traits de personnalité, par essence plus stables que des symptômes. En Europe, les pionniers de ce que Pichot appellera la *psychopathologie quantitative* (Pichot 1972) sont, précisément, Pierre Pichot (1954, 1961) et Max Hamilton (1959, 1960, 1961).

Après deux décennies d'applications souvent sauvages d'échelles d'évaluation psychiatriques trop rarement validées, les psychiatres restent passionnels dans leur attitude face à ces instruments, les inconditionnels scotomisant les contraintes de la méthode, les rebelles niant leur nécessité et leur intérêt. Ces échelles d'évaluation du changement en psychopathologie n'ont d'ailleurs toujours pas trouvé droit de cité dans les traités et dictionnaires classiques de psychiatrie : elle ont été "oubliées" (sic) de la dernière édition du Porot (1984) ; elles n'ont leur place dans aucun traité français - si ce n'est, en ce qui concerne les échelles d'anxiété et de dépression, dans la "Psychiatrie de l'adulte" de Guelfi (1985) - ni, jusqu'ici, dans l'Encyclopédie médico-chirurgicale (Guelfi & Bobon, en préparation). Elles sont mentionnées du bout de la plume dans la nouvelle édition du "Lexikon der Psychiatrie" édité par Christian Müller (1986). Quant à la dernière édition du "Comprehensive Textbook of Psychiatry" américain, elle ne consacre aucune section aux échelles d'évaluation dans le chapitre 6 sur les méthodes quantitatives et expérimentales en psychiatrie tandis que, dans la section 12.6 sur le testing psychologique de la personnalité, le psychologue Arthur C. Carr (1985) semble confondre les critères diagnostiques qualitatifs (type DSM-III), les tests de personnalité proprement dits (type MMPI) et les échelles d'évaluation de la psychopathologie.

Le lecteur intéressé trouvera en revanche des références sur les échelles les plus connues dans le rapport du congrès de psychiatrie et de neurologie de langue française de Charleroi (Guelfi et al. 1978), dans le chapitre du "Précis pratique de psychiatrie" canadien sur les tests psychologiques (Viguié 1984), dans le premier ouvrage français sur les échelles d'évaluation (Cottraux et al. 1985) et dans un récent répertoire analytique des échelles sensibles au changement (Bobon 1987), sans oublier, en langue anglaise, le catalogue - malheureusement épuisé - de la "Early Clinical Drug Evaluation Unit" ou ECDEU (Guy 1976) et, en langue alle-

mande, le catalogue d'un groupement de laboratoires pharmaceutiques allemands et suisses, le "Collegium Internationale Psychiatriæ Scalarum" (1981).

Quant au lecteur intéressé par des réflexions récentes sur le principe de la quantification de la psychopathologie, il sera avisé de consulter un numéro des *Acta Psychiatrica Belgica* consacré à des "Mélanges à Jean Bobon" sur la recherche en psychiatrie, en particulier les contributions de Hamilton, de Lehmann et de Luccioni (1983).

Nous allons, quant à nous, passer en revue des pièges de la quantification de la psychopathologie que nous avons rencontré depuis quinze ans dans nos lectures et dans les travaux de l'École de Liège sur les applications des échelles d'évaluation aux essais thérapeutiques.

### Auto- ou hétéro-évaluation ?

Dans les essais thérapeutiques, de nombreux travaux convergent pour indiquer que les échelles remplies par les thérapeutes, y compris par le personnel infirmier, sont plus sensibles au changement que les échelles d'auto-évaluation.

Cependant, l'auto-évaluation garde toute sa valeur dans les études sur de grands échantillons, en particulier en épidémiologie (par exemple dans le dépistage de tendances dépressives et/ou anxieuses chez des malades somatiques). En outre, l'auto-évaluation trouve depuis peu un nouveau champ d'application dans le domaine du testing informatisé par micro-ordinateur ; dès 1977, Pichot et al. confirmaient des travaux américains selon lesquels certains patients sont plus en confiance devant l'écran d'un ordinateur que devant le regard d'un psychiatre ou d'un psychologue, sans compter les aspects d'objectivité et de mode dont bénéficie actuellement l'informatique. Un dernier avantage de la passation sur micro-ordinateur : elle permet un calcul instantané du profil factoriel d'un malade par comparaison à un échantillon de référence.

### Echelles globales, spécifiques ou multifactorielles ?

Différents rapports ont conclu à la valeur des évaluations globales du type "Visual Analogue Scales" (VAS) - que nous avons proposé d'appeler "lignes de Scott" en hommage à leur inventeur - ou "Clinical Global Impressions" (CGI). L'évaluation globale, dans ce qu'elle a de synthétique et d'intuitif, vaut ce que vaut le clinicien qui procède à cette évaluation. Ceci est d'ailleurs vrai pour toutes les échelles psychiatriques, celui qui la remplit (qu'il s'agisse d'un thérapeute ou du malade lui-même) faisant partie intégrante de la mesure ; que la qualité du cotateur conditionne la qualité de la mesure peut paraître un rappel trivial mais chacun sait que cette condition n'est pas toujours respectée. Or, les

analyses de données ne pourront pas faire un Stravivarius d'un mauvais bois : le danger de GIGO ("Garbage in, garbage out") guette chaque étude basée sur des échelles mal validées ou mal remplies.

Certaines échelles, comme la "Brief Psychiatric Rating Scale" (BPRS) ou l'échelle syndromique du groupe francophone de l'AMDP (l'AMDP-SY), proposent une évaluation globale d'un certain nombre de fonctions psychopathologiques ; elles ont donc les qualités et les faiblesses de toute évaluation globale. Nous reviendrons plus loin sur une faiblesse propre à la BPRS : le nombre excessif de degrés des items.

Quant aux échelles spécifiques (de la dépression, de l'anxiété etc), elles sont le plus souvent préférées aux échelles multifactorielles ou "comprehensive" en raison de leur rapidité de passation. Si nous avons introduit dès 1974 une échelle multifactorielle dans tous nos essais pilotes, c'est parce que nous étions - et restons - convaincu qu'à ce stade de l'évaluation de spectre d'activité clinique d'un nouveau psychotrope, le fait de se contenter d'échelles spécifiques met des œillères à la recherche : en n'utilisant qu'une échelle de dépression, on ne voit pas si le médicament à l'étude améliore ou aggrave l'anxiété ; lorsque la note de dépression s'effondre, on ignore si le malade est guéri ou a viré en manie.

Nous avons souvent affirmé que réaliser des études de phase II en double aveugle, éventuellement multicentriques sinon multinationales, avec tout le poids éthique, pratique et financier que l'on sait, en réduisant toute l'information clinique à la seule note totale d'une échelle spécifique, c'est se condamner à ne voir que la partie émergée de l'iceberg. Le choix entre échelle spécifique ou multifactorielle ne doit pas être guidé par une logique économique mais par une logique scientifique : en règle générale, une échelle multifactorielle est indispensable en phase II alors qu'une échelle spécifique est suffisante en phase III.

### Problème de validité

Trop d'expérimentateurs ou de laboratoires pharmaceutiques ont tendance à choisir telle ou telle échelle parce qu'elle est d'usage courant, facilitant les comparaisons avec d'autres publications ou avec des travaux antérieurs du même expérimentateur ou du même laboratoire ; c'est ainsi que la sacro-sainte "banque de données" empêche souvent l'apport d'échelles nouvelles, mieux validées ou mieux adaptées à ce que l'on veut mesurer. Un compromis raisonnable consiste à inclure à la fois une échelle de référence et, si nécessaire, une nouvelle échelle.

Lorsqu'on utilise une échelle traduite d'une autre langue, il faudrait toujours vérifier si la *traduction*

n'a pas altéré le sens initial de l'instrument. On a ainsi montré que la traduction française des assertions du "Minnesota Multiphasic Personality Inventory" (MMPI) déforme parfois la signification des assertions américaines d'origine au point d'invalider l'imputation de la réponse du patient à telle ou telle échelle du MMPI. Lorsque nous avons traduit la "Befindlichkeits-Skala" ou BfS de D. von Zerssen en français avec l'aide de trois autres spécialistes supposés excellents bilingues, nous avons notablement déformé le sens des adjectifs allemands dans 16 % des cas ; c'est en suivant les recommandations de l'Organisation Mondiale de la Santé que nous avons découvert ces distorsions, faisant procéder à une *rétroversion* ("back-translation") du français en allemand par un cinquième bilingue réunissant deux conditions nécessaires : avoir pour langue maternelle la langue d'origine de l'échelle et, bien sûr, n'avoir jamais lu la version originale de cette échelle. Combien de traductions ont-elles été vérifiées par ce procédé ?

Lorsqu'on touche à des concepts psychopathologiques, une traduction erronée de l'intitulé d'un item peut conduire à ce qu'on ne mesure pas en français la même chose que dans la langue d'origine. Dans certains cas, la rétroversion est insuffisante ou inadaptée et doit céder la place à des cotations vidéo par des participants bilingues. Un exemple extrait de nos travaux d'adaptation des items psychopathologiques du Système AMDP en français est celui de la pensée "umständlich", que nous avions initialement traduit par "digressive" alors que le concept allemand sous-jacent à l'adjectif "umständlich" correspond à l'adjectif français "circonstancié" (cf Bobon 1985, p. 176), décalage que seules des cotations vidéo franco-allemandes ont permis de déceler.

Le libellé des questions ou items doit être univoque. L'exemple du MMPI est classique : dans l'espoir d'éviter des biais systématiques de réponse par le sujet, par exemple systématiquement "vrai", les psychologues de l'Université du Minnesota ont conçu des assertions formulées tantôt positivement, tantôt négativement, de même que des phrases sous forme de double négation. On connaît les conséquences de ces théories de laboratoire : les malades, ne comprenant pas, répondent souvent de façon aléatoire. Il en est de même des lignes de Scott : en bonne théorie psychométrique, il nous est arrivé d'alterner des lignes où le pathologique était au pôle droit et des lignes où le pathologique était au pôle gauche. Conséquence : les malades mais aussi les thérapeutes, conservant leurs automatismes, cochaient systématiquement du même côté ce qu'ils considéraient comme pathologique, d'où des réponses non valides. Nous avons donc acquis la conviction que tant le mode de réponse que la formulation des items ou questions doivent être simples et systématiques sous peine de diminuer la validité des notes obtenues.

On oublie souvent qu'afin de savoir ce qu'une échelle est censée mesurer, il faut savoir sur quel échantillon elle a été validée. Ainsi, l'échelle d'anxiété de Cattell, validée sur une majorité de "petits" anxieux, n'est pas adaptée à la mesure de l'anxiété pathologique alors que l'échelle d'anxiété de Hamilton, destinée à des malades psychiatriques hospitalisés, n'est pas adaptée à la mesure de l'anxiété de cas légers.

L'échelle d'anxiété de Hamilton (HAMA) va nous permettre d'illustrer un autre aspect des problèmes de validité. L'HAMA a été construite sur la base d'un concept tiré de la clinique : celui d'une participation importante des répercussions somatiques de l'anxiété à la mesure de cette fonction, d'où l'inclusion a priori de 50 % d'items somatiques par Hamilton dans son échelle. Or, nous avons montré par analyse factorielle, sur un échantillon de 493 malades neuropsychiatriques de tous diagnostics, que le facteur Anxiété ne comporte qu'un seul item somatique sur dix-sept ou, encore, un item somatique sur les 47 items somatiques évalués (Bobon et al. 1985). Autrement dit, dans les limites de notre échantillon, les plaintes somatiques de malades psychiatriques ne sont des répercussions de leur anxiété que pour 1 item sur 17, soit 16 % des cas : nous sommes loin des 50 % de Hamilton. Ceci pourrait expliquer pourquoi, sur un autre échantillon, des malades considérés par évaluation globale comme peu ou pas anxieux obtiennent à l'HAMA une note très supérieure à la note qu'ils obtiennent lors de la même passation à l'échelle tirée du facteur d'anxiété de l'AMDP, l'AMDP-AT (Bobon et al. 1985).

La contamination de l'anxiété par la dépression et vice versa est un autre exemple de la validité discutable de certaines échelles ; chacun sait par exemple qu'un anxiolytique peut faire significativement baisser la note à l'échelle de dépression de Hamilton (HAMD) en raison de la présence, dans cette échelle, d'un nombre important d'items communs aux deux pathologies. Lorsqu'on souhaite une mesure indépendante de la dépression ou de l'anxiété, il est sans doute préférable de porter son choix sur des échelles extraites d'analyses factorielles basées sur l'orthogonalité entre ces deux fonctions, c'est-à-dire sur le calcul de facteurs ayant entre eux une corrélation nulle grâce aux saturations des items retenus.

La prise en compte des saturations dans le calcul d'une note totale nous amène à évoquer un vieux débat : celui des *pondérations*. Si la puissance et la rapidité de calcul des ordinateurs modernes permet aujourd'hui de transformer aisément des notes d'items brutes en notes pondérées en fonction de la saturation de ces items dans le facteur considéré, pareille pondération continue à rester pratiquement inappliquée. Pourquoi ? Sans doute parce que, si la nature des items d'un facteur varie peu d'un échantillon à l'autre, le poids respectif de ces

items dans le facteur peut varier considérablement, empêchant la généralisation d'une pondération. Nous pensons cependant que ce problème des pondérations devrait faire l'objet de nouvelles études car l'utilisation des notes brutes d'items, telle qu'elle est actuellement pratiquée, affaiblit évidemment le caractère orthogonal, indépendant, des facteurs.

Un dernier aspect de la validité qui nous a frappé lors de nos lectures et recherches est celui des *notes-seuils*. Certaines échelles sont appliquées depuis plus de vingt ans sans que personne n'ait, semble-t-il, jamais calculé statistiquement la note discriminant le plus significativement les sujets, par exemple, anxieux des non-anxieux. Lorsque nous avons posé ce problème au statisticien de notre Service, Rémy von Frenckell, il a, comme en d'autres occasions, trouvé un test statistique classique adapté à cette application clinique inédite, en l'occurrence le test t de Student. C'est ainsi qu'on été calculées des notes-seuils de dépression pour l'HAMD (von Frenckell & Lottin 1982) et pour l'AMDP-AT (Bobon et al. 1985).

### Problèmes de fidélité

Les problèmes de fidélité des mesures, surtout celui de la fidélité intercotateurs, conditionnant la validité mais plus encore la sensibilité de ces mesures, ont bien plus polarisé l'attention des utilisateurs que les problèmes de validité. Nous serons donc plus brefs sur ce chapitre.

Nous avons déjà mentionné le problème de l'intitulé des items et de leur définition comme le libellé des questions et le mode de réponse.

Le choix des items eux-mêmes influence l'accord entre cotateurs. C'est ainsi qu'on a cru dans les années cinquante que la fidélité intercotateurs serait meilleure pour des items comportementaux, observables, que pour des items sur le vécu des malades, d'où des échelles comme celles de Wittenborn ou de Lorr. Dès 1968, Shepherd et al. montraient qu'il n'en est rien. Ces mêmes auteurs ont également insisté sur l'influence de l'a priori diagnostique sur la cotation des items : face au même enregistrement vidéo, les cotateurs pensant à une schizophrénie cochent l'item "flattening of affects" tandis que les cotateurs pensant à une mélancolie cochent "depressive apathy" ; nous avons retrouvé dans nos travaux cette sélection des items correspondant à un a priori diagnostique (Bobon 1985).

Dans les cas de concepts complexes et/ou d'échelles longues, un *entretien semi-structuré* s'est révélé, selon notre expérience, indispensable à une passation complète, sans déformations ni oublis, surtout dans le cas de cotateurs insuffisamment rodés à la psychopathologie et/ou à l'instrument.

La recherche du *nombre optimal de degrés de réponse* aurait pu bénéficier des observations de docimolo-

gie dans l'enseignement. En pratique, les auteurs d'échelles psychiatriques n'ont pratiquement jamais vérifié l'assimilation implicite que nous faisons entre valeurs ordinales et valeurs numériques en transformant des catégories "léger-moyen-fort" en "1-2-3". Par exemple R. von Frenckell (1979) a montré par la technique de reparamétrisation ou "optimal scaling" que 8 degrés de gravité comme Overall le préconise dans ses échelles (échelle de dépression, BPRS) dépassent les capacités de discrimination du jugement humain. La gradation qui paraît le meilleur compromis entre la fidélité intercotateurs (absent/présent) et la sensibilité (plusieurs degrés) est celle qui va de 0 à 4 ; dans la version allemande des échelles AMDP, que nous avons initialement maintenue en français, nous avons noté que la gradation de 0 à 3 n'est pas en réalité assimilable à quatre degrés mais à un degré de normalité et à trois degrés de pathologie ("léger-moyen-fort"), d'où la *tendance centrale* à abuser de la catégorie "moyen".

L'inclusion de *critères de gradation* accroît certes la fidélité intercotateurs mais, dans le cas d'items complexes comme les items correspondant à des concepts psychopathologiques comme les phobies, la spécification des critères de cotation restreint la diversité des situations cliniques, d'où le risque d'une moins bonne validité. Nous pensons cependant que, malgré ce risque, la délimitation précise des critères de gravité constitue un élément indispensable à une évaluation fidèle.

Pareille délimitation contribue d'ailleurs à résoudre l'hésitation du cotateur quant au poids qu'il doit donner respectivement à l'*intensité*, à la *durée* et à la *fréquence* des symptômes. Un problème connexe est celui de la *période évaluée* (par exemple une semaine ou un mois) : nous avons souvent observé que l'absence d'instructions quant à la période de temps couverte par une cotation est une source de désaccords aussi importante que facile à prévenir.

Il n'est pas certain que, malgré toutes les précautions méthodologiques prises, le problème des différences de *seuils culturels* de perception du caractère pathologique ou non d'un symptôme puisse être éliminé.

Ce problème est connu depuis longtemps mais nous en avons trouvé un nouvel exemple dans nos travaux, celui d'une patiente marseillaise en manie délirante, Fernande, dont les items d'humeur (euphorie, surestimation de soi, idées délirantes de grandeur, logorrhée etc) ont été évalués comme moins pathologiques par les cotateurs marseillais que par des cotateurs "nordiques" (Bobon 1985).

### Problèmes de sensibilité

La sensibilité d'une échelle est largement liée à la fidélité de l'évaluation : cf le nombre de degrés des

items, les critères de gradation etc discutés plus haut. Voyons brièvement certains aspects de la sensibilité qui n'ont pas encore été abordés.

Le plus classique est celui de l'inclusion, dans une même échelle, de *traits* (de personnalité le plus souvent) et d'*états* momentanés (des symptômes, par exemple).

Une question plus originale est la suivante : la technique de cotation en *temps aveugle* ("time-blind") diminue-t-elle la sensibilité de l'évaluation du changement en supprimant la référence temporelle ? On sait que la cotation en temps aveugle d'enregistrements vidéo avant et après traitement vise à réduire le risque, conscient ou inconscient, que le thérapeute cote "high" avant inclusion dans un essai et "low" en fin d'essai. Nous n'avons connaissance que d'une seule publication (Mormont et al. 1984) qui montre, sur un échantillon de 9 cas, que la cotation des enregistrements avant et après traitement d'un malade en temps aveugle et lors de séances espacées d'au moins une semaine ne diminue pas la sensibilité de la mesure par rapport à des cotations qui facilitent a priori un jugement comparatif en supprimant le "time gap", c'est-à-dire par l'évaluation des deux enregistrements d'un même malade lors d'une même séance.

Enfin, une erreur malheureusement fréquente consiste à élaborer des échelles destinées à être sensibles au changement sur la base d'analyses factorielles de cas indépendants, c'est-à-dire d'analyses ne comportant qu'une seule passation par malade.

Or, un facteur, constitué d'items corrélés entre eux, peut ne pas être sensible au changement.

Aussi certains s'orientent-ils aujourd'hui vers la construction de *facteurs de changement* (Lapierre & von Frenckell 1983, Digonnet et al. 1987), au risque que ces facteurs ne soient pas généralisables à d'autres médicaments mais cette voie de recherche nous paraît prometteuse.

## Conclusion

La psychiatrie de demain passe par la recherche d'aujourd'hui. Les échelles d'évaluation soulèvent de nombreux problèmes mais elles sont devenues indispensables. Nous pensons avec Bech et al. (1982) que le temps d'un moratoire est venu dans le domaine des échelles d'évaluation de la psychopathologie, cessant d'en construire de nouvelles avant d'avoir démontré par des études poussées la validité et les lacunes des instruments existants.

*Remerciements : nous tenons à profiter de ce bilan d'un parcours de vie pour remercier les personnes qui nous ont donné le goût de la rigueur et de la recherche (les Professeurs Jean Bobon et Jean Lecomte, de l'Université de Liège) et celles qui ont guidé nos réflexions et nos travaux dans le domaine de la psychopathologie quantitative (les Professeurs Hamilton, Pichot et von Zerssen, ainsi que notre ami Remy von Frenckell).*

## BIBLIOGRAPHIE / REFERENCES

AITKEN R.C.B., SOUTHWELL N., WILMSHURST C.C.  
- Chloral betaïne: an effective hypnotic.  
*Clin-Trials J.* 1965; 2: 65-70.

BECH P., GASTPAR M., MOROZOV P. - Clinical assessment scales for biological psychiatry to be used in WHO studies,  
In : *Rapport à la réunion de Lucknow, Inde, Octobre 1981.* Organisation Mondiale de la Santé, Genève, 1982.

BOBON D. - Contribution de l'échelle AMDP à la psychopathologie quantitative. Thèse.  
*Acta Psychiatr Belg* 1985; 85: 1-249.

BOBON D. - Psychopathologie quantitative et mesure du changement,  
In: MENDLEWICZ J. ed.: *Manuel de Psychiatrie biologique*, Masson, Paris, 1987.

BOBON D., von FRENCKELL R., TROISFONTAINES B., MORMONT C., PELLET J. - Construction et validation préliminaire d'une échelle d'anxiété extraite de l'AMDP francophone, l'AMDP-AT.  
*Encéphale* 1985 ; 11 : 107-111.

BONNET C. - *Contemplation de la nature*, tome 1. Marc-Michel Rey, Amsterdam, 1770.

CARR A.C. - Psychological testing of personality.  
In: KAPLAN H.I. & SADOCK B.J. eds.: *Comprehensive textbook of psychiatry*, 4th ed., Williams and Wilkins, Baltimore, 1985.

CATTELL J. McK. - Mental tests and measurements.  
*Mind* 1890; 15: 373-380.

COLLEGIUM INTERNATIONALE PSYCHIATRIAE SCALARUM ed. : *Internationale Skalen für Psychiatrie*, 3<sup>e</sup> ed. Beltz Test, Weinheim, 1986.

COTTRAUX J., BOUVARD M., LEGERON P. - *Méthodes et échelles d'évaluation des comportements*. EAP, Issy-les-Moulineaux, 1985.

DELAY J., PICHOT P., PERSE J. - *Méthodes psychométriques en clinique*. Masson, Paris, 2<sup>e</sup> ed., 1966.

DIGONNET C., BOYER P., PUECH A.J., von FRENCKELL R., WEIL J-S. - Les échelles AMDP dans l'évaluation de psychotropes en phase II.  
*Acta Psychiatr Belg* 1987 ; 87 : 245-252.

GALTON F. - Psychometric experiments.  
*Brain* 1879; 2: 149-162.

- GUELF J.D. - *Psychiatrie de l'adulte*, Ellipses, Paris, 1985.
- GUELF J.D., DREYFUS J-F., PULL C.B. - Les essais thérapeutiques en psychiatrie : méthodologie, éthique et législation. In : *76<sup>e</sup> Congr Psychiatr neurol franç* Masson, Paris, 1978.
- GUELF J.D., BOBON D. - Echelles d'évaluation et critères de diagnostic en psychiatrie. *Encycl Med Chir, Psychiatrie*, Paris, en préparation.
- GUY W. - *ECDEU Assessment Manual for Psychopharmacology*. 3rd ed. National Institute of Mental Health, Rockville, 1976.
- HAMILTON M. - The assessment of anxiety states by rating. *Br J Med Psychol* 1959; 32: 50-55.
- HAMILTON M. - A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; 23: 56-62.
- HAMILTON M. - *Lectures on the methodology of clinical research*. E. & S. Livingstone, Edinburgh, 1961.
- HAMILTON M. - Us et abus des échelles d'évaluation et des statistiques en psychopharmacologie. *Acta Psychiatr Belg* 1983 ; 83 : 155-162.
- LAPIERRE Y.D., von FRENCKELL R. - AMDP psychopathology factors in chronic schizophrenia: a clinical trial of two long-acting neuroleptics. In: BOBON D. et al. eds. : *The AMDP-System in pharmacopsychiatry*, 193-203, Karger, Basel, 1983.
- LEHMANN H.E. - De Binet et Wundt aux mesures neuropsychologiques et aux échelles de comportement. *Acta Psychiatr Belg* 1983 ; 83 : 135-146.
- LUCCIONI H. - Quelques mots sur la quantification de la clinique dans la recherche en psychiatrie. *Acta Psychiatr Belg* 1983 ; 83 : 163-168.
- MORMONT C., von FRENCKELL R., LOTTIN T., MORMONT Y., TROISFONTAINES B., BOBON D. - Influences de variables temporelles (time-blind, time gap) sur l'évaluation quantitative de la psychopathologie. *Encéphale* 1984 ; 10 : 3-7.
- MULLER C. ed.- *Lexikon der Psychiatrie*. 2<sup>e</sup> éd. Springer, Berlin, 1986.
- PICHOT P. - L'évolution de l'emploi et des recherches sur les tests mentaux aux États-Unis. I. Les tests de personnalité. *Rev Psychol Appl* 1954 ; 4 : 317-340.
- PICHOT P. - Les tests de personnalité en psychiatrie. In : COSSA P. ed. : *54<sup>e</sup> Congr méd alién neurol franç*, Masson, Paris, 1956.
- PICHOT P. - L'évaluation des modifications du comportement humain sous l'influence des drogues psychotropes. In : ROTHLIN E. ed. : *Neuropsychopharmacology*, II, 304-318, Elsevier, Amsterdam, 1961.
- PICHOT P. - Psychopathologie quantitative. *Neurologie-Psychiatrie* (Sandoz) 1972 ; 3 : 3-9.
- PICHOT P., DREYFUS J.F., DREYFUS M., GUELF J., SOMOGYI M., SEMELIN J. - Rating of an auto questionnaire presented on-line by a computer to psychiatric patients. In: BERTELLI A. et al. eds.: *Evaluation of new drugs in clinical psychopharmacology*. Prous J.R., Barcelona, 1977.
- POROT A. ed. - *Manuel alphabétique de psychiatrie*, 6<sup>e</sup> éd., Presses universitaires de France, Paris, 1984.
- SCOTT W.D. - Do you want to know what others think about you? *American Magazine* 1920; Nov: 44-47.
- SHEPHERD M., BROOKE E.M., COOPER J.E., LIN T. - An experimental approach to psychiatric diagnosis. An international study. *Acta Psychiatr Scand* 1968; Suppl. 201.
- VIGUIE F. - Les tests psychologiques, In : DUGUAY R. & ELLENBERGER H.F. eds. : *Précis pratique de psychiatrie*, 2<sup>e</sup> éd - St-Hyacinthe, Paris et Edisem, Maloigne, 1984.
- von FRENCKELL R. - *Première application de la méthode de reparamétrisation Prinqual en psychopathologie quantitative*. Thèse, Clin. Psychiatr. Univ., Liège, 1979.
- von FRENCKELL R., LOTTIN T. - Validation d'un seuil de la dépression : l'échelle d'Hamilton. *Encéphale* 1982 ; 8 : 349-354.
- WITTENBORN J.R. - A new procedure for evaluating mental hospital patients. *J Consult Clin Psychol* 1950; 14: 500-501.
- WITTENBORN J.R. - *Wittenborn Psychiatric Rating Scales*. Psychological Corp., New York, 1955.

# ÉTATS DÉPRESSIFS RÉOUVRIR «L'ANGLE DE VIE»

P 75

# PROTHIADEN 75

## PROTHIADEN 75

**Forme et présentation:** comprimés enrobés (rouge): boîte de 20.  
**Composition:** dosulépine (DCI) chlorhydrate ou chlorhydrate de dihydro-6-11 (diméthylamino propylidène) dibenzo (b, e) thiépine. Par comprimé enrobé: 75mg, par boîte de 20: 1.500g. **Propriétés:** antidépresseur à action anticholinergique périphérique et centrale. Lève l'inhibition dépressive; modifie l'humeur dépressive dans le sens normo-thymique; tend à apaiser l'anxiété dépressive; ses manifestations somatiques. **Indications:** Les syndromes dépressifs de toute nature. Ce dosage est également adapté aux traitements d'entretien. **Contre-indications:** l'association de Prothiaden 75 avec les inhibiteurs de la mono-amine-oxydase avec le Prothiaden 75 est formellement contre indiquée; il est nécessaire de laisser un intervalle libre de 15 jours entre une cure par I.M.A. et un traitement par le Prothiaden 75. Glaucome à angle fermé. Risque de rétention urinaire chez les sujets porteurs d'une hypertrophie prostatique. **Précautions:** les précautions d'emploi du Prothiaden 75 sont celles des anti-dépresseurs tricycliques. En particulier: conduite d'engin, la consommation d'alcool, les sujets épileptiques ou ayant des antécédents épileptiques, en présence d'une hypertension et chez les malades présentant des antécédents cardiaques, grossesse et période d'allaitement. **Effets indésirables:** les effets indésirables sont liés aux effets périphériques et centraux des anti-dépresseurs tricycliques: manifestations de type atropinique, tremblements, troubles du sommeil. **Mode d'emploi et posologie:** dans les dépressions légères, la posologie habituelle est de 75mg, soit un comprimé enrobé le soir. Dans les dépressions d'intensité moyenne, la posologie peut être portée à 150mg en une à deux prises dès l'amélioration des symptômes. La posologie pourra être réduite à 75mg par jour. En milieu psychiatrique spécialisé, la posologie peut atteindre 300mg/jour. Coût moyen de traitement journalier: 2,00F. Tableau AMM, 325009.7 (1981) - Mis sur le marché en 1984. **Prix:** 39,90F - S.H.P. (20 comprimés enrobés). Remb. Séc. Soc. à 70% - Collect.

Laboratoires Boots-Dacour.  
49 rue de Bitche  
B.P.66 - 92404 Courbevoie Cedex  
Tél. (1) 47.88.51.44 - Télex: 610.67