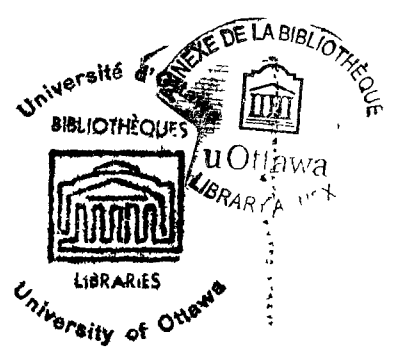


PREDICTIONS ACTUAIRE ET CLINIQUE DU COMPORTEMENT
AGRESSIF CHEZ DES SCHIZOPHRENES AU MOYEN DU HAND TEST

par William Rodrigus

Thèse présentée à la Faculté de Psychologie
de l'Université d'Ottawa en vue de satisfaire
partiellement aux exigences de la Maîtrise
en Psychologie.



Ottawa, Canada, 1968

UMI Number: EC56250

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EC56250
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

RECONNAISSANCE

Cette thèse fut préparée sous la direction du professeur Gilles Chagnon, M.Ps., de la Faculté de Psychologie de l'Université d'Ottawa.

Les suggestions apportées par Laurent Isabelle, Ph.D., ont avantageusement modifié la revue des écrits. On remercie les personnes qui ont participé à l'expérience en tant que juge.

CURRICULUM STUDIORUM

William Rodrigue naquit à Duparquet, province de Québec, le 25 janvier 1943. Il obtint son B.A. du Collège de Rouyn, province de Québec, en 1964.

TABLE DES MATIERES

Chapitres	pages
INTRODUCTION	viii
I.- RETROSPECTIVE ET PERSPECTIVE	1
1. Valeur du test comme instrument de prédiction	1
2. Valeur du test comme instrument de prédiction de l'agression	5
3. Considérations théoriques et hypothèse générale	11
II.- LES MODES DE PREDICTION ACTUAINE ET CLINIQUE . . .	16
III.- LE PLAN D'ETUDE, Première partie	22
1. Principes de sélection du matériel projectif	22
2. Choix des études de Drummond et de Medvedeff	25
3. L'emploi de protocoles déjà existants	26
4. Description des études de Drummond et de Medvedeff	28
5. Evaluation critique de ces deux études	30
IV.- LE PLAN D'ETUDE, Deuxième partie	36
1. Représentations de la méthode clinique dans les écrits	37
2. Exigences de la prédiction clinique avec le <u>Hand Test</u>	41
3. Présente représentation de la méthode clinique	43
V.- LE PLAN D'ETUDE, Troisième partie	47
1. Un échantillon stratifié de protocoles	47
2. Reproduction et adaptation des protocoles	48
3. Modalités de la tâche imposée aux juges	52
4. Définitions opératoires et hypothèses nulles	54
5. Opérations statistiques	58

TABLE DES MATIERES

v

Chapitres	pages
VI.- LA PRESENTATION DES RESULTATS	60
1. Constance des corrections actuaires	61
2. Validité des corrections actuaires	68
3. Constance de la méthode clinique	71
4. L'efficacité prédictive des trois groupes de juges comparés	75
5. Comparaison des méthodes actuaire supé- rieure et clinique moyenne	78
6. Comparaison des méthodes actuaire et cli- nique supérieures	81
7. Comparaison des méthodes actuaire et cli- nique moyennes	83
8. Comparaison des méthodes actuaire moyenne et clinique supérieure	83
9. La prédiction clinique dans les études de Drummond et de Medvedeff	85
VII.- L'ANALYSE DES RESULTATS	88
1. Evaluation de la constance	88
2. Evaluation de la validité	89
3. Comparaisons des modes de prédiction	91
4. La portée des résultats	94
RESUME ET CONCLUSIONS	97
BIBLIOGRAPHIE	100
Appendices	
1. <u>INSTRUCTIONS</u>	104
2. <u>SAMPLE ANSWER SHEET</u>	107
3. <u>REPRODUCTION OF THE 9 HAND TEST CARDS</u>	109
4. CARACTERISTIQUES INDIVIDUELLES DES JUGES	111
5. ABSTRACT OF <u>Prédictions actuaire et clinique du comportement agressif chez des schizo- phrènes au moyen du Hand Test</u>	118
6. SOMMAIRE DE <u>Prédictions actuaire et clinique du comportement agressif chez des schizo- phrènes au moyen du Hand Test</u>	121

LISTE DES TABLEAUX

Tableaux	pages
<p>I.- Représentation des groupes agressifs (A) et non agressifs (NA) des études de Medvedeff et de Drummond dans l'échantillon actuel quant à la distribution des nombres et des sexes, en termes de proportions</p>	49
<p>II.- Représentation des groupes agressifs (A) et non agressifs (NA) des études de Medvedeff et de Drummond dans l'échantillon actuel quant aux succès respectifs obtenus par le prédiseur d'agression (<u>A.O.S.</u>) et l'index de non interaction (<u>WITH</u>), en termes de proportions</p>	50
<p>III.- Le pourcentage d'accord des présentes corrections actuaires entr'elles et avec les corrections actuaires de Drummond et de Medvedeff</p>	62
<p>IV.- Le pourcentage d'accord entre les diverses corrections actuaires du prédiseur d'agression (<u>A.O.S.</u>) et de l'index de non interaction (<u>WITH</u>)</p>	64
<p>V.- La constance interne des correcteurs de la présente expérience sur le prédiseur d'agression (<u>A.O.S.</u>) et sur l'index de non interaction (<u>WITH</u>) en termes de pourcentages</p>	65
<p>VI.- Les valeurs de chi deux correspondant à l'accord des diverses corrections actuaires du prédiseur d'agression (<u>A.O.S.</u>) et de l'index de non interaction (<u>WITH</u>)</p>	67
<p>VII.- La signification des diverses corrections actuaires du prédiseur d'agression (<u>A.O.S.</u>) et de l'index de non interaction (<u>WITH</u>) dans les protocoles tirés des études de Drummond et de Medvedeff, en valeurs de chi deux</p>	69
<p>VIII.- La signification des différences d'efficacité prédictive entre les diverses corrections actuaires du prédiseur d'agression (<u>A.O.S.</u>) et de l'index de non interaction (<u>WITH</u>) des protocoles tirés de l'étude de Medvedeff, en valeurs de chi deux</p>	72

Tableaux	pages
IX.- La signification des différences d'efficacité prédictive entre les diverses corrections actuelles du prédiscur d'agression (A.O.S.) et de l'index de non interaction (<u>WITH</u>) des protocoles tirés de l'étude de Drummond, en valeurs de chi-deux	73
X.- Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiscur d'agression (A.O.S.) et l'index de non interaction (<u>WITH</u>) dans les corrections actuelles originales et par le groupe total de juges, et la signification des différences en valeurs de "t"	80
XI.- Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiscur d'agression (A.O.S.) et l'index de non interaction (<u>WITH</u>) dans les corrections actuelles originales et par le groupe de juges supérieurs, et la signification des différences en valeurs de "t"	82
XII.- Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiscur d'agression (A.O.S.) et l'index de non interaction (<u>WITH</u>) dans les diverses corrections actuelles et par le groupe total de juges, et la signification des différences en valeurs de "t"	84
XIII.- Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiscur d'agression (A.O.S.) et l'index de non interaction (<u>WITH</u>) dans les diverses corrections actuelles et par le groupe de juges supérieurs, et la signification des différences en valeurs de "t"	86

INTRODUCTION

La majorité des tests projectifs ont pour but de faciliter l'étude de la personnalité en découvrant les tendances, les besoins, les conflits et les modes d'adaptation de l'individu. On leur accorde cet avantage particulier de révéler des réalités psychologiques peu accessibles à une approche plus directe: soit que la personne étudiée se montre défensive, soit qu'elle ignore les informations recherchées. Comme ces techniques permettent en général une évaluation globale de la personnalité, elle connaissent un usage répandu dans les différents domaines de la psychopathologie. Un de ces domaines a trait à l'évaluation du contrôle émotif.

Créé en 1960 par Edwin E. Wagner, le Hand Test¹ a pour but premier d'évaluer les tendances à l'action, c'est à dire les tendances perceptuomotrices et leur hiérarchie dans leurs interactions avec le monde extérieur. De là, il n'y avait qu'un pas à franchir pour que soit abordé le problème de l'évaluation du contrôle émotif. Le présent travail a pour objectif d'évaluer les mérites relatifs des approches actuaire et clinique comme moyens de prédiction du

¹ E. E. Wagner, E. Bricklin et Z. A. Piotrowski, The Hand Test, A New Projective Test with Special Reference to the Prediction of Overt Aggressive Behavior, Springfield, Illinois, Charles C. Thomas, 1962, x-100 p.

comportement agressif avec le Hand Test.

Le premier chapitre présentera une revue critique des études de validation sur le test. Les considérations découlant de cette revue permettront de formuler l'hypothèse générale suivante: l'emploi d'une approche clinique, représentée par un groupe de juges, peut améliorer l'efficacité prédictive du système de correction présentement en usage. Le fait de substituer une approche clinique au système de correction lie notre travail à la controverse qui sépare les adeptes de l'une et l'autre approches, actuaire et clinique. Le deuxième chapitre sera donc consacré à une revue sélective des études traitant des mérites relatifs à chaque approche.

Le plan d'étude comprend trois chapitres. On procédera d'abord au choix du matériel projectif qui va permettre la comparaison statistique des deux approches. Il s'agira ensuite de représenter adéquatement le mode de prédiction clinique. Enfin, la procédure employée, les définitions opératoires, les hypothèses de recherche et les opérations statistiques seront décrites. Aux sixième et septième chapitres les résultats de l'expérience seront présentés et analysés.

Le dernier chapitre consiste en un résumé de la thèse et se terminera par quelques suggestions à l'intention des chercheurs qui vont pousser plus loin le travail de validation du test. On a cru à propos d'insérer en appendice les informations que les juges avaient en main, ainsi que la

INTRODUCTION

x

liste des caractéristiques individuelles des juges.

CHAPITRE PREMIER

RETROSPECTIVE ET PERSPECTIVE

La première partie de ce chapitre a pour but de présenter une revue rapide et une appréciation globale des recherches expérimentales rapportées dans les écrits sur le Hand Test, en donnant une importance particulière aux travaux qui ont traité de la prédiction de l'agression. Cette recension mettra en lumière la discordance générale dans les résultats obtenus et conduira à l'étude du développement du système de correction et des améliorations apportées au cours de sa validation progressive. On avancera enfin quelques hypothèses susceptibles d'orienter les efforts de recherche vers une évaluation scientifique plus poussée du test comme instrument de prédiction de l'agression.

1. Valeur du test comme instrument de prédiction.

Depuis la parution du Hand Test en 1960, on compte un total de dix-neuf études ayant trait à sa validation. De ces études, 9 ont obtenu des résultats positifs et sur les dix autres, 4 seulement ont obtenu des résultats à peine marginaux. De plus, tous les résultats positifs furent rapportés par Wagner et ses disciples: c'est dire qu'aucune des recherches indépendantes revues n'a répondu à l'attente de son auteur, quelle que fut l'hypothèse avancée.

La variation des hypothèses de recherche des études pourrait expliquer une part des divergences observées. Toutefois, deux études avaient pour but spécifique la contrevalidation de recherches déjà faites par Wagner et ses collaborateurs.

Il s'agit en premier lieu de l'étude de Huberman¹ tentant de valider celle de Wagner et Copper².

Wagner and Copper (1963) reported high discriminative power for one score (the ACTivity score) of the Wagner Hand Test between (unskilled and semi-skilled) workers at Goodwill Industries, Akron, Ohio who had been rated "Satisfactory" or "Unsatisfactory" by their superiors.

An attempt to crossvalidate this finding, in a three-group design, in a large western Plywood mill failed to reproduce the expected relationship; neither the ACT, nor any of the other scores derived from this test showed any consistent trend in terms of post-dicting activity level, or general acceptability of the 18 Ss involved in this study.

¹ J. Huberman, A Failure of the Wagner Hand Test to Discriminate among Workers Rated High, Average and Low on Activity Level and General Acceptability, dans Journal of Projective Techniques and Personality Assessment, vol. 28, n° 3, livraison de septembre 1964, p. 280.

² E. E. Wagner et J. Copper, Differentiation of Satisfactory and Unsatisfactory Employees at Goodwill Industries with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 27, n° 3, livraison de septembre 1963, p. 354-356.

Il s'agit encore de l'étude de Drummond³ qui n'a pas su différencier des schizophrènes agressifs de schizophrènes non agressifs en dépit de la différence largement significative rapportée trois ans plus tôt, en 1963, par Wagner et Medvedeff⁴. Ces deux études seront étudiées plus loin en détail.

Puisque la variété des hypothèses de recherches n'explique pas la discordance des résultats obtenus, du moins pour les deux cas mentionnés, on se voit obligé de recourir à d'autres sources d'information si l'on veut poursuivre notre évaluation critique du Hand Test comme instrument de prédiction.

Tout d'abord, certaines critiques sont adressées aux conclusions tirées par les auteurs qui ont obtenu les résultats escomptés. Entr'autres, la critique un peu amère de Shaw, qui insiste sur la nécessité d'une distinction entre la validité de prédiction et la validité concurrente.

³ F. Drummond, A Failure in the Discrimination of Aggressive Behavior of Undifferentiated Schizophrenics with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 30, n° 3, livraison de juin 1966, p. 275.

⁴ E. E. Wagner et E. Medvedeff, Differentiation of Aggressive Behavior of Institutionalized Schizophrenics with the Hand Test, dans Journal of Projective Techniques, vol. 27, n° 1, livraison de mars 1963, p. 111-113.

A cursory perusal of the research done so far reveals that the claims for prediction of overt behavior are totally unjustified and are, in fact, in contradiction with the test parameters enumerated in the manual for administration.

[...] the assumption of predictive validity is totally unsupported by research and appears to be based upon what has been termed "common sense validity [...] the product of confident ignorance⁵."

D'autres estiment qu'il est illégitime de conclure à l'efficacité prédictive du test au sein de populations non connues puisque la majorité des chercheurs ont employé la méthode dite des groupes connus.

Cette généralisation excessive n'enlève pas à notre avis la valeur et l'utilité de la méthode des groupes connus, surtout au moment des premières tentatives de validation d'un test. Nous admettons cependant qu'ultimement il sera nécessaire de recourir à des études longitudinales, en vue d'évaluer la validité prédictive du Hand Test, et à des études faites au sein de populations non connues, en vue de scruter si son pouvoir de discrimination peut soutenir une application plus générale.

L'emploi d'une approche plus rigoureuse à ce stage-ci de développement du Hand Test pourrait en effet conduire à un rejet prématuré d'un instrument potentiellement valable.

⁵ D. L. Shaw et J. D. Linden, A Critique of the Hand Test, dans Educational and Psychological Measurement, vol. 24, n° 2, livraison d'été 1964, p. 283-284.

Voilà pourquoi on doit être conscient de son apparition assez récente, comme le suggère la revue faite par Mills⁶ et, avec Gleser, insister davantage sur ses aspects encore non explorés :

The Hand Test appears to have possibilities for development as a quantitative multidimensional clinical test relevant to the overt behavior of individuals. However, its psychometric characteristics are essentially unknown so that a great deal more developmental work is needed in order to shape the technique into a measuring instrument rather than a springboard for metaphorical interpretations⁷.

2. Valeur du test comme instrument de prédiction de l'agression.

Les études donnant une importance au moins secondaire à la prédiction de l'agression, au nombre de dix, reflètent bien la discordance générale que l'on vient de noter.

⁶ D. H. Mills, The Research Use of Projective Techniques: A 17 Year Survey, dans Journal of Projective Techniques and Personality Assessment, vol. 29, n° 4, livraison de décembre 1965, p. 513.

⁷ G. C. Gleser, faisant la revue du Hand Test dans la section Character and Personality (p. 141-540) de The Sixth Mental Measurements Yearbook, O. K. Buros, éditeur, Highland Park, N.-J., The Gryphon Press, 1965, p. 437.

A nouveau Wagner et ses disciples, Hawkins⁸, Wetsel et Shapiro⁹, Medvedeff¹⁰, Hodge et Schreiner¹¹, ont remporté un succès complet à quatre reprises. Chaque étude investiguait la valeur du prédicteur d'agression, le Acting Out Score, ou de scores en faisant partie. Les résultats obtenus sont tous significatifs au niveau de .01 ou .001, sauf pour le score AFV (affection) dans l'étude de Hodge et Schreiner où le seuil de signification atteint est de .025.

La diversité des échantillons employés dans ces études, des échantillons de normaux, de schizophrènes et de délinquants, portait à croire que le test pourrait sans crainte affronter de nouvelles épreuves. Ce fut loin d'être le cas, si l'on juge par les recherches suivantes qui n'ont pourtant pas, à première vue du moins, soumis le test à des conditions expérimentales plus exigeantes.

8 E. E. Wagner et R. Hawkins, Differentiation of Assaultive Delinquents with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 28, n° 3, livraison de sept. 1964, p. 363-365.

9 E. E. Wagner, H. Wetsel et R. J. Shapiro, Prediction of Recidivism among Juvenile Delinquents with the Hand Test, dans Journal of Projective Technique and Personality Assessment, vol. 31, n° 4, livraison de sept. 1967, p. 69-72.

10 E. E. Wagner et E. Medvedeff, op. cit., p. 111-113.

11 E. E. Wagner, J. Hodge et F. Schreiner, Hypnotic Validation of Two Hand Test Scoring Categories, dans Journal of Projective Techniques and Personality Assessment, vol. 30, n° 4, livraison d'août 1966, p. 365-366.

Des six chercheurs indépendants, seuls Oswald et Foucar-Egyed rapportent des conclusions légèrement favorables au test. Oswald rapporte que des cinq catégories composant le prédicteur d'agression, que l'en explicitera en détail plus loin, seulement deux ont atteint le seuil requis de signification (.05):

Normative data obtained on the Hand Test from 114 secondary school boys, 52 institutionalized male delinquents and 26 institutionalized female delinquents were compared. Significant differences were in the expected direction for certain scores. Comparison with the "equivalent" American samples given by Wagner shows marked differences. Looseness of definition of certain scoring categories affected interscorer reliability¹².

Et bien que le chi deux obtenu pour le prédicteur d'agression fut significatif au niveau de .01, l'auteur remarque certaines différences:

The differences between Wagner's and the present writer's groups are obvious and are most marked between the Acting Out Natives. It is doubtful whether these differences are a result of age differences between the groups; for no age trends were shown either in an unpublished study by G. Gloss of a group of 205 children ranging from 7 to 15 years or within the scoring categories of the "normal" boys in our sample who ranged from 13 to 18 years¹³.

¹² M. O. Oswald et P. Loftus, A Normative and Comparative Study of the Hand Test with Normal and Delinquent Children, dans Journal of Projective Techniques and Personality Assessment, vol. 31, n° 4, livraison de septembre 1967, p. 62.

¹³ Idem, ibidem, p. 64.

La précision des hypothèses expérimentales, le choix des opérations statistiques et la triple correction indépendante des protocoles suggèrent que l'expérience fut menée avec soin. Le procédé d'échantillonnage paraît aussi adéquat. Les adolescents normaux furent choisis au hasard dans une population de près de 1200 élèves du niveau secondaire d'une école technique. Le quotient intellectuel moyen dans l'échantillon était de 101, avec un écart type de 11.4. Aucun de ces élèves n'avait de dossier judiciaire. Les protocoles des cinquante-deux garçons délinquents furent pris au hasard parmi les résultats de test du nombre de délinquents qui avaient, sur une période de trois ans, subi une évaluation par le service de Psychologie attaché au département du Bien Etre Social. Tous ces adolescents étaient déjà parus en cours juvénile pour offense criminelle. Une procédure semblable fut employée pour l'échantillon de filles délinquantes.

Foucar-Egyed avait un an plus tôt, dans une thèse non publiée, étudié les relations entre les tendances à l'agression et l'âge chronologique chez des enfants de trois groupes d'âges différents (six ans, neuf ans et treize ans). Ses résultats vont à l'encontre de l'explication hypothétique avancée par Oswald voulant que l'âge ne soit pas une variable à contrôler.

Although there is not a significant difference between each and every age group, it may be concluded from the results that aggression, as measured by the AOR of The Hand Test, does decrease with chronological age. Also, it may be predicted that aggression, as measured by the Aggression responses of The Hand Test, would decrease even with the increase in the total number of responses with chronological age¹⁴.

Le test, comme instrument de mesure, pourrait bien manquer de constance. On note que Foucar-Egyed rapporte une correspondance de 90,5% entre deux corrections indépendantes alors qu'Oswald n'a obtenu que 73%. Il devient alors difficile d'attribuer une valeur certaine à cette généralisation de Foucar-Egyed:

The Spearman Rank-difference Correlation figure of .548 is significant at the .01 level of confidence. Thus, it may be concluded that the AOR is a reliable measure in this study with the subjects used. Assuming this score to be representative of the reliability of other test scores, it may be concluded further that the test itself is reliable¹⁵.

Il semble plus prudent d'adopter les vues de Gleser¹⁶ qui estime que l'inclusion de quelques cartes, comme stimuli additionnels, élèverait de façon appréciable la constance du test. On sait que le nombre moyen de réponses au test chez une

¹⁴ G. Foucar-Egyed, Overt Aggression Expressed between the Ages of 6-9-13, as Defined by Scores on The Hand Test, thèse de Maîtrise présentée à l'École de Psychologie et d'Éducation, Ottawa, 1966, p. 68.

¹⁵ Idem, ibidem, p. 63.

¹⁶ G. C. Gleser, op. cit., p. 436-438.

personne "normale" n'est que de 11.3 ($Q_3 - Q_1 = 4.1$)¹⁷.

Au terme de cette revue, il devient difficile d'accepter sans réserves les premières conclusions rapportées par Wagner.

The acting out score successfully differentiated overtly aggressive adult groups from non aggressive groups. Not only did the score differentiate acting out from non acting out groups, but also many non acting out populations among themselves. Prison inmates and hospital acting out cases earned significantly higher acting out scores than did indigents, non acting out hospital cases, and normals. In addition, the indigents, non acting out hospital cases, and normals were all differentiated among themselves¹⁸.

Il semble au contraire que le test n'a pas encore atteint le premier but qu'on lui avait fixé:

The first goal was to determine whether the test responses would differentiate among groups on whom information was possessed concerning aggressive acting out tendencies¹⁹.

Pour cette raison, le présent travail estime nécessaire d'orienter les efforts de recherche vers une évaluation scientifique plus poussée du test comme instrument de prédiction de l'agression. Nous allons maintenant avancer quelques hypothèses susceptibles d'expliquer la discordance dans les

17 E. E. Wagner, The Hand Test, Manual for Administration, Scoring, and Interpretation, Akron, Ohio, Mark James, 1962, p. 20.

18 E. E. Wagner, B. Bricklin et Z. A. Piotrowski, The Hand Test, A New Projective Test with Special Reference to the Prediction of Overt Aggressive Behavior, Springfield, Illinois, Charles C. Thomas, 1962, p. 85.

19 Idem, ibidem, p. 4.

résultats obtenus.

3. Considérations théoriques et hypothèse générale.

On observe tout d'abord que la presque totalité des études a trait à la validation du système de correction présentement en usage. Ce système pourrait bien être à l'origine d'un bon nombre des divergences observées. Si c'est le cas, reprendre le travail déjà fait en tentant une autre validation indépendante serait voué au départ à l'échec. Ne serait-il pas avantageux d'emprunter une approche différente qui permettrait d'éviter à cette difficulté majeure? Afin de déceler le bien-fondé de cette affirmation pour l'instant gratuite, une revue critique du système de correction s'avère nécessaire. Une importance particulière sera accordée aux altérations qu'a subi le système au cours de son développement en tant qu'instrument de prédiction de l'agression.

La monographie, première publication majeure sur le test, résumait le système de correction en onze catégories ou classes de réponses. Le prédicteur d'agression se calculait ainsi:

The acting out score is obtained by subtracting the sum of responses in the classes of Fear, Affection, Communication, and Dependence from the sum of responses in the classes of Aggression and Directs, i.e., $\text{Sum (Aggression + Directs)} - \text{Sum (Fear + Affection + Communication + Dependence)}$ ²⁰.

²⁰ Idem, ibidem, p. 7.

Au cours de la même année, le manuel du test paraissait, introduisant quatre nouvelles différentiations. On notait aussi le retrait du score Fear du prédiseur d'agression. Par contre, les classes de réponses, maintenant au nombre de quinze, étaient regroupées en quatre divisions majeures. Une de ces divisions, appelées le Withdrawal Score, que l'on traduit ici par index de non interaction, allait permettre la création d'un second prédiseur d'agression dans les populations de schizophrènes. Cet index de non interaction se compose des réponses descriptives, des réponses bizarres et des refus de répondre.

Cette dernière version du système de correction a fait l'objet d'étude de dix-sept des dix-neuf recherches mentionnées au début; il ne semble donc pas que les améliorations apportées aient ajouté à la capacité prédictive du test. D'autres altérations ou additions substantielles sont probablement nécessaires. Cette supposition gagnerait sans doute à être soumise à une étude systématique. C'est également l'opinion de Fisher qui affirme en 1967, après une revue critique du développement des techniques projectives:

Increasingly analyses have appeared in the literature which indicate that apparent contradiction in results merely reflect different phases or levels of a response system²¹.

²¹ S. Fisher, Projectives Methodologies, dans Annual Review of Psychology, vol. 18, 1967, p. 176.

En admettant que le système de correction soit erroné ou insuffisamment complexe, il devient nécessaire de revoir la méthode qui a permis sa création.

The problems in validating a test such as this may be schematized as follows. First there is the (conceptual) task of grouping the responses into categories which would seemingly have meaning and utility for the psychologist. The next step is to specify the behavioral empirical equivalents of the groups or categories of responses. These first two steps are usually inter-dependant, since the formalized groupings are generally made with an eye to the empirical behaviors which it is believed are reflected by the categories. The next step is the confrontation of the first two steps (i.e., the comparison of the formal categories with the behavioral equivalents) within the confines of a theory of scientific confirmation or disconfirmation²².

Il semble que le travail de conceptualisation auquel les auteurs font allusion relève proprement de l'approche hypothético-déductive. L'incapacité relative du système de correction à permettre la réussite des tentatives de validation pourrait résulter d'erreurs ou d'omissions commises au moment de cette phase critique de développement du test. Meehl se permet d'affirmer:

²² E. E. Wagner, B. Bricklin et E. A. Piotrowski, *op. cit.*, p. 49.

"How much information is intrinsically contained in the test behavior?" We are beginning to collect some evidences with regard to the M.M.P.I. that nobody gets all of the information out and I would assume that to be equally true of less structured device²³.

Comparé au M.M.P.I., le test de la Main sort à peine de ses langes. Une réévaluation des informations qui n'ont pas été incorporées au système de correction paraît d'autant plus impérieuse. Si l'on juge à partir de l'importance accordée à l'évaluation qualitative dans la monographie et le manuel du test, on constate que les auteurs eux-mêmes sont pleinement conscients de cette lacune.

Clearly not all aspects of Hand Test interpretation can be readily quantified. Qualitative evaluations are crucial secondary methods of interpretation and can be of extreme value to the perspicacious investigator, although it must be equally clear that qualitative analyses make greater demands on the flexibility and clinical insight of the examiner²⁴.

Outre les difficultés inhérentes à la quantification et l'incorporation de ces données cliniques, on doit envisager au préalable un problème primordial: est-ce que ces informations non quantifiées, et possiblement non quantifiables, peuvent ajouter à la valeur du test comme instrument de prédiction de l'agression? C'est là le but du présent travail:

²³ P. E. Meehl, Structured and Projective Tests: Some Common Problems in Validation, dans Journal of Projective Techniques, vol. 23, n° 3, livraison de sept. 1959, p. 268-272.

²⁴ E. E. Wagner, The Hand Test, Manual for Administration, Scoring, and Interpretation, p. 31.

comparer l'efficacité prédictive d'une approche clinique, représentée par un groupe de juges, à la méthode actuariale, représentée par le système de correction présentement en usage dans le test de la Main.

Une telle position, toutefois, s'oppose de front à la controverse très actuelle quant aux mérites des deux approches, actuariale et clinique. Le prochain chapitre présentera une revue sélective des écrits dans ce domaine.

CHAPITRE II

LES MODES DE PREDICTION ACTUAIRE ET CLINIQUE

Le chapitre précédent a démontré à partir d'une appréciation globale des écrits sur le Hand Test que le système de correction, même après amélioration, s'avère insatisfaisant. Nous avons alors suggéré l'emploi d'une approche clinique afin de vérifier si le système actuariaire n'a pas omis des données pertinentes qui pourraient être repérées par une appréciation clinique. Il se pourrait en effet qu'il y ait eu des erreurs d'émission ou de jugement dans le processus hypothético-déductif qui a conduit à la formulation des catégories ou classes de réponses composant le système de correction. Il se pourrait encore que la prédiction de l'agression nécessite une flexibilité que le système de correction ne possède pas mais qui serait un attribut du jugement clinique.

A cause de la nature spécifique du Hand Test, un instrument projectif qui emploie un système de correction actuariaire, le fait de substituer une approche clinique à l'approche actuariaire lie le présent travail de près à la controverse qui sépare les adeptes de l'une et l'autre approches. Une revue sélective des écrits dans ce domaine devient alors nécessaire. Notons au préalable que l'objectif poursuivi ici n'est pas de prouver la supériorité d'une

approche sur l'autre mais bien de contribuer à la validation du test. Notre intention est de prendre connaissance des écrits les plus récents afin qu'il soit possible de juger si les résultats obtenus à date justifient l'emploi d'une approche clinique pour le but que l'on poursuit.

Comme notre but consiste à évaluer l'écart qui sépare le potentiel d'information recélé dans le matériel projectif et le montant d'information contenu dans le système de correction, les juges représentant l'approche clinique seront constamment obligés de critiquer et d'évaluer les scores obtenus par l'approche actuariale, c'est à dire par le système de correction du test. Cette condition expérimentale s'oppose directement à cet énoncé de Meehl qui, loin de croire aux vertus de l'approche clinique, va jusqu'à dire:

Mostly we will use our heads because there just isn't any formula, but suppose we have a formula, and a case comes along in which it disagrees with our heads? Shall we then use our heads? I would say, yes - provided the psychological situation is as clear as a broken leg; otherwise, very, very seldom.

En quelque sorte à la tête du mouvement en faveur de l'approche actuariale, Meehl compte à sa suite un bon nombre

1 Paul K. Meehl, When Shall We Use Our Heads Instead of the Formula?, dans Journal of Counseling Psychology, vol. 4, n° 4, livraison d'hiver 1957, p. 273.

de disciples: Hathaway², Kleinmuntz³, Watley⁴, et plusieurs autres qu'il serait trop long d'énumérer.

Cette citation de Meehl allait susciter beaucoup d'intérêt. Entr'autres, elle fut reprise par Sawyer en 1966 qui lui répond:

...rather we should try to find the optimal combination of actuarially controlled method and sensitive clinical judgement for any particular enterprise⁵.

Parmi ses partisans importants, on compte Schaeffer⁶ qui minimise les différences entre les tests objectifs et projectifs; Holt⁷ qui souligne la nécessité de trouver un

2 S. R. Hathaway, Clinical Intuition and Inferential Accuracy, dans Journal of Personality, vol. 24, n° 3, livraison de mars 1956, p. 223-250.

3 Benjamin Kleinmuntz, Sign and Seer: Another Example, dans Journal of Abnormal Psychology, Vol. 72, n° 2, livraison d'avril 1967, p. 163-165.

4 Donovan J. Watley, Do Counselors Know When to Use Their Heads Instead of the Formula, National Merit Scholarship Corp. Research Reports, vol. 3, n° 1, 1967 12 p.

5 J. Sawyer, Measurement and Prediction, Clinical and Statistical, dans The Psychological Bulletin, vol. 66, n° 3, 1966, p. 183.

6 R. W. Schaeffer, Clinical Psychologist's ability to use the Draw-A-Person Test, dans Journal of Consulting Psychology, vol. 28, n° 4, livraison d'août 1964, p. 383.

7 R. Holt, Clinical and Statistical Prediction: A Re-Formulation and Some New Data, dans Journal of Abnormal and Social Psychology, vol. 56, n° 1, livraison de janvier 1958, p. 1-12.

juste milieu entre les deux approches; McArthur⁸ qui, pour sa part, croit que le jugement clinique a droit de cité puisqu'il s'apparente de près au développement d'une théorie, que nul statisticien ne saurait substituer par une équation mathématique.

Au lieu de prendre part à la controverse comme telle, un certain nombre de chercheurs préfèrent opter pour une position mitoyenne, une synthèse qui pourrait se formuler en ces termes; pour Cooke⁹, les deux approches valent parce qu'elles peuvent conduire à une prédiction actuariale supérieure; pour Forer¹⁰, le clinicien peut apprendre à améliorer son pouvoir de prédiction en utilisant les succès obtenus par la méthode actuariale; d'autres chercheurs, enfin, comme Gordon¹¹ et Johnston¹², estiment que les deux modes

⁸ C. C. McArthur, Analysing the Clinical Process, dans Journal of Counseling Psychology, vol. 1, livraison d'hiver 1954, p. 203-207.

⁹ J. K. Cooke, Clinicians' Decisions as a Basis for Deriving Actuarial Formulae, dans Journal of Clinical Psychology, vol. 23, n° 2, livraison d'avril 1967, p. 232-233.

¹⁰ B. R. Forer, Sentence Completions, dans The Prediction of Overt Behavior Through the Use of Projective Techniques, Springfield, Ill., Charles C. Thomas, 1960, p. 6-17.

¹¹ Leonard V. Gordon, Clinical, Psychometric and Work Sample Approaches in the Prediction of Success in Peace Corps Training, dans Journal of Applied Psychology, vol. 51, n° 2, livraison d'avril 1967, p. 111-119.

¹² Roy Johnston et B. McNeal, Statistical versus Clinical Prediction: Length of Neuropsychiatric Hospital Stay, dans Journal of Abnormal Psychology, vol. 72, n° 4, livraison d'août 1967, p. 335-340.

se valent et qu'il est préférable de s'en tenir à des critères pratiques lorsqu'il s'agit de choisir une approche particulière dans un contexte donné.

Les écrits qui viennent d'être mentionnés sont relativement récents et confirment le caractère actuel de la controverse. C'est dire qu'il peut être encore avantageux de recourir à un mode de prédiction clinique.

Des tendances énumérées, celle représentée par Cooke, qui croit que les deux approches valent parce qu'elles peuvent conduire à une prédiction actuariale plus satisfaisante, semble davantage compatible avec l'objectif du présent travail. Par le fait même, on s'identifie également à la position de Sawyer, qui estime préférable de trouver le juste milieu entre les deux approches pour une tâche donnée. Cette position, toutefois, pourrait bien ne plus valoir à un stage ultérieur de développement du système de correction du test de la Main. Mais à ce stage-ci de son développement, nous lui donnons raison et répondons à Meehl qu'il est nécessaire et possiblement fructueux d'employer son jugement lorsque la formule actuariale en usage se montre inadéquate. Loin d'être en voie d'extinction, la méthode clinique intéresse un nombre grandissant de chercheurs sérieux. C'est du moins ce que suggèrent les conclusions tirées par Fisher dans sa revue sur l'avancement des techniques projectives:

A growing array of investigators continues to believe that imaginative responses to ink blots and pictures can be sensibly codified and used to make worth-while predictions. Contrary to the grim judgments of observers who have chided "projective" researchers for their "faith" and unwillingness to be realistic, the present reviewer detects increasing objectivity and fair-mindedness in their investigative strategies. Many serious and well designed studies have been published which sincerely seek to learn both the potentialities and the limitations of specific projective tests and their associated scoring systems¹³.

¹³ S. Fisher, Projective Methodologies, dans Annual Review of Psychology, Vol. 18, 1967, p. 184.

CHAPITRE III

LE PLAN D'ETUDE

Le présent travail a pour but de comparer l'efficacité prédictive d'une approche clinique, représentée par un groupe de juges, à la méthode actuariale, représentée par le système de correction présentement en usage dans le Hand Test. Le chapitre précédent a confirmé la légitimité de ce but en démontrant qu'il est nécessaire et possiblement fructueux de recourir à une approche clinique lorsque le système actuariale en usage se montre inadéquat.

Ce chapitre-ci va d'abord traiter des principes de sélection du matériel projectif. C'est à la lumière de ces principes que les études de Drummond et de Medvedeff ont été choisies comme source du matériel projectif. Il est alors à propos d'examiner les implications inhérentes à l'emploi de protocoles déjà existants. On procédera ensuite à une description et à une évaluation critique de ces deux études afin de juger si elles ont fait preuve d'une rigueur scientifique suffisante pour servir de point de départ au présent travail.

1. Principes de sélection du matériel projectif.

Wagner rapporte qu'il est plus aisé d'opérer un diagnostic aveugle dans une population anormale que dans

une population normale.

In general, the more pathological the responses the less one needs to refer to other data in assessing the subject's functioning efficiency and the greater the probability of making correct "blind diagnoses": a psychotic cannot function well regardless of his environmental supports¹.

Wetsel, dans la même ligne, soutient qu'une interprétation subjective ou clinique des scores numériques obtenus devient surtout nécessaire quand il s'agit de sujets anormaux.

Further research should [...] take into account not only the extent of aggressive proclivities but also the intensity of psychopathology which may generally depress action tendencies or put them beyond conscious control. These tendencies may, however, express themselves sporadically and intermittently under conditions of lowered conscious control. The cut-off points should therefore be used with caution, especially when dealing with protocols where marked psychopathology is present. [...] Prediction for the individual case is therefore hazardous and should be undertaken only in the context of the entire protocol and, of course, all the other information at the disposal of the diagnostician².

Pour ces raisons, la présente recherche utilisera des protocoles de schizophrènes afin de donner tous les avantages possibles à la méthode clinique.

¹ Edwin E. Wagner, The Hand Test, Manual for Administration, Scoring, and Interpretation, Akron, Ohio, Mark James, 1962, p. 2.

² H. Wetsel, R. J. Shapiro, et E. E. Wagner, Prediction of Recidivism among Juvenile Delinquents with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 31, n° 4, livraison de sept. 1967, p. 71.

Une étude faite par Girardin en 1964 n'a pas réussi à démontrer une différence significative entre des schizophrènes simples et des schizophrènes paranoïaques quant à leur comportement sur le Hand Test, sauf pour le nombre de réponses appartenant à la catégorie BIZ (réponses bizarres). L'auteur attache cependant peu de valeur à cette découverte:

Out of 23 studies involving direct comparisons between the paranoid and simple schizophrenics on various aspects of the Hand-Test, mathematical differences were always present except once; however, only one of the differences was significant statistically. This was on the Bizarre sub-category of scoring where the difference was found to be significant beyond the .01 level of confidence. It is to be said that due to the low frequency of these Bizarre responses, the obtained significant difference should not be given too much consideration³.

Comme ces réponses se prêtent facilement à une interprétation clinique, suivant la pensée de Wetsel, on ne voit pas d'objection à utiliser des schizophrènes non différenciés. De plus, les réponses bizarres appartiennent au WITE, qui est employé comme un second prédicteur d'agression chez les schizophrènes. Ce serait alors injuste qu'une catégorie de schizophrènes obtiennent un score plus élevé sur ce prédicteur, indépendamment de ses tendances à l'agression. La souplesse du jugement clinique pourrait peut être d'une part, obvier à cette difficulté du système actuel et, d'autre part, conduire à une différenciation supplémentaire capable de rétablir

³ N. B. Girardin, The Use of the Hand Test for Differentiating Simple and Paranoid Schizophrenics, Thèse de maîtrise présentée à l'École de Psychologie et d'Éducation, Ottawa, 1964, p. 47.

ce déséquilibre.

2. Choix des études de Drummond et de Medvedeff.

Les études de Drummond⁴ et de Medvedeff⁵ avaient pour but précis d'évaluer la capacité du Hand Test à prédire les tendances à l'agression chez des schizophrènes non différenciés. Suivant les principes de sélection du matériel projectif qui viennent d'être énoncés, il semble nécessaire d'accorder une attention spéciale à ces deux études.

En 1963, Medvedeff et Wagner réussirent à différencier des schizophrènes agressifs de schizophrènes non agressifs en utilisant le prédiscur d'agression et l'index de non interaction. Ils rapportent des résultats significatifs aux niveaux de .01 et de .001, respectivement.

En 1966, Drummond reproduit exactement la même expérience en vue de valider en Angleterre les résultats obtenus aux Etats-Unis et n'obtint aucune différence significative entre les deux groupes, bien que son critère d'agression ait

⁴ F. Drummond, A Failure in the Discrimination of Aggressive Behavior of Undifferentiated Schizophrenics with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 30, n° 3, livraison de juin 1966, p. 276-279.

⁵ E. E. Wagner et E. Medvedeff, Differentiation of Aggressive Behavior of Institutionalized Schizophrenics with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 27, n° 1, livraison de mars 1963, p. 111-113.

été davantage exigeant que celui de Medvedeff. Ces résultats contrastants semblent bien refléter le degré précaire de développement atteint par le système de correction présentement en usage.

La présente recherche se propose d'employer les protocoles utilisés par les études de Drummond et de Medvedeff afin d'évaluer le mérite de l'approche clinique dans une tâche où l'approche actuariaire opère à des niveaux de succès significativement différents l'un de l'autre. L'acceptation dans notre recherche des données brutes provenant d'études indépendantes peut toutefois recéler un inconvénient d'importance, celui de priver le clinicien d'une observation sur place du patient. Ce point mérite considération.

3. L'emploi de protocoles déjà existants.

Quelle est donc la valeur d'une prédiction clinique faite à partir de réponses projectives écrites plutôt qu'à partir d'une observation sur place du sujet par le clinicien, qui administre alors lui-même le test?

Grebstein⁶ et Hoffman croient que la situation de testing, étant particulière, ne permet pas de reproduire par

⁶ L. Grebstein, Relative Accuracy of Actuarial Prediction, Experienced Clinicians, and Graduate Students in a Clinical Judgement Task, dans Journal of Consulting Psychology, vol. 27, n° 2, livraison d'avril 1963, p. 127-132.

la suite les mêmes données de base.

...unstructured clinical judgement situations nonetheless make the contribution of such information experimentally impossible to assess. The problem here under investigation would, as a result, cease to be a scientific problem altogether⁷.

Sawyer⁸ estime au contraire que les méthodes clinique et actuaires devraient être représentées à leur meilleur et que tous les indices susceptibles de permettre une prédiction plus réussie peuvent s'employer. Soumettant ce problème à une étude scientifique, Grigg ne rapporte cependant aucune différence significative entre les deux conditions:

...reading interview statements was associated with as accurate predictions as occurred after face to face interaction with the client. The script condition, with total absence of voice cues, was associated with the most accurate predictions with two of the three measuring instruments used⁹.

Comme certains vont jusqu'à refuser à la méthode clinique les informations exclues de la méthode actuaires, ce qui signifierait dans notre cas le rejet des réponses verbales au test, et comme d'autres insistent sur la liberté nécessaire

7 P. J. Hoffman, The Paramorphic Representation of Clinical Judgement, dans Psychological Bulletin, vol. 57, n° 2, 1960, p. 118.

8 J. Sawyer, Measurement and Prediction, Clinical and Statistical, dans Psychological Bulletin, vol. 66, n° 3, 1966, p. 178-200.

9 A. E. Grigg, Experience of Clinicians, and Speech Characteristics, and Statements of Clients as Variables in Clinical Judgement, dans Journal of Consulting Psychology, vol. 22, n° 4, livraison d'août 1958, p. 319.

au clinicien dans le recrutement de ses indices, la position prise ici est motivée par l'objectif du présent travail: celui de contribuer à la validation progressive du test. C'est dire que les juges représentant l'approche clinique auront accès à tous les indices susceptibles d'améliorer leurs prédictions à condition, toutefois, que ces indices soient communicables, propres à être reproduits et, si possible, propres à être quantifiés. Il semble alors logique d'exclure la situation de testing.

On procédera maintenant à une description plus complète et à une évaluation critique des études de Drummond et de Medvedeff afin de juger si elles ont fait preuve d'une rigueur scientifique suffisante pour justifier l'acceptation de leurs données brutes dans le présent travail.

4. Description des études de Drummond et de Medvedeff.

Dans l'étude de Medvedeff, les sujets étaient des schizophrènes non différenciés d'un hôpital psychiatrique de l'Ohio. Des psychiatres et des infirmières classifièrent les patients en deux groupes, agressif et non agressif, d'après les critères suivants, pour le groupe agressif: présence d'actes ou de paroles hostiles, destruction de propriété, tentative de fuite, refus de coopération dans les activités proposées ou dans les soins apportés; et pour le groupe non agressif: attitude généralement coopérative ou passive,

acceptation volontaire des tâches proposées et absence d'actes destructifs. Les cas douteux furent exclus des deux groupes. Chaque groupe compte trente-cinq patients: 21 hommes et 14 femmes dans le groupe agressif; 18 hommes et 17 femmes dans le groupe non agressif. Un chi deux de 7.06, corrigé par la formule de Yates, fut obtenu pour le prédiseur d'agression. Ce résultat s'avère significatif au niveau de .01. Un chi deux corrigé de 12.87, significatif au niveau de .001, fut obtenu pour le score de non interaction.

L'étude de Drummond voulant contrevalider celle de Medvedeff, la sélection des sujets et le choix des critères furent les mêmes dans les deux études. Toutefois, pour une meilleure représentation du groupe agressif, Drummond n'y inclut que ceux qui avaient fait preuve de violence physique ou de destruction de propriété depuis au moins cinq ans. Les critères d'acceptation pour le groupe non agressif ne furent en rien altérés. Les sujets présentant des symptômes organiques furent rejetés, entr'autres: épileptiques, lobotomisés, patients traités par chocs électriques. La différenciation des deux groupes fut faite à partir de notes déjà inscrites aux dossiers des patients. Drummond ne fait aucune mention de l'examineur et du correcteur. On sait cependant qu'il(s) ignorait(aient) tout des sujets, sauf que ceux-ci avaient été clairement diagnostiqués schizophrènes. Les protocoles contenant des réponses difficiles à corriger et qui

auraient pu altérer le prédiscur d'agression ou le score de non interaction furent éliminés.

Avant de procéder à l'analyse des résultats, Drummond vérifia statistiquement la possibilité d'une différence entre les deux groupes aux points de vue de l'âge et de la durée d'hospitalisation: aucune différence significative ne fut obtenue au niveau de .05. La distribution des sexes était comme suit: 19 hommes et 17 femmes dans le groupe agressif; 15 hommes et 15 femmes dans le groupe non agressif. Un chi deux corrigé de .04, non significatif à .05, fut obtenu pour le prédiscur d'agression; et un chi deux corrigé de .15, non significatif à .05, fut obtenue pour l'index de non interaction. Drummond conclut que ses résultats, loin de valider ceux de Medvedeff, font au contraire ressortir une ressemblance marquée entre ses groupes agressif et non agressif quant à leur comportement sur le prédiscur d'agression et l'index de non interaction du test.

5. Evaluation critique de ces deux études.

Un premier point qui mérite réflexion est bien celui de la définition du terme "comportement agressif". Le choix des critères qui ont servi à la différenciation des groupes agressifs et non agressifs chez l'un et l'autre auteurs fut fait à la lumière de la définition suivante, proposée par Wagner en 1962.

"Acting-out" was defined for our purposes as a subject's behaving in such a way as to bring him to the attention of others (police, court, school authorities, guidance clinics psychiatrists, etc.) as a result of overt aggressive behavior. The "acting-out score" is not devised to predict specific motor acts, but rather a tendency to act out in an aggressive way of any kind. Thus, our concept of acting out includes such activities as: instigating legal proceedings against others for delusional paranoid reasons; robbing or swindling others; throwing objects with the intent of hitting (as a wife or husband in marital quarrels); the school child's picking fights with his peers or teachers; violent destruction of bedding or furniture by a disturbed mental patient, etc.¹⁰.

Les critères d'agression choisis par Drummond et Medvedeff semblent bien refléter le genre de tendance à l'agression que le Hand Test est censé mesurer.

Deux autres variables qui auraient pu induire en erreur sont l'âge et l'intelligence.

Tous les sujets employés dans les deux études étaient âgés d'au moins vingt ans. Cette précaution paraît justifiée si l'on se rapporte à l'étude de Foucar-Egyed¹¹ en 1966 et l'appréciation de Wagner en 1962 qui suggèrent une relation significative entre l'âge et le développement des tendances sociales positives et des tendances sociales négatives de la

¹⁰ E. E. Wagner, B. Bricklin et G. A. Piotrowsky, The Hand Test, A New Projective Test with Special Reference to the Prediction of Overt Aggressive Behavior, Springfield, Illinois, Charles C. Thomas, 1962, p. 4-5.

¹¹ C. Foucar-Egyed, Overt Aggression Expressed between the Ages of 6-9-13, as Defined by Scores on the Hand Test, Thèse de maîtrise présentée à l'École de Psychologie et d'Éducation, Ottawa, 1966, 73 p.

personnalité. Bien que ce ne fut pas confirmé par Oswald¹² en 1967, qui croit que cette relation disparaît après l'âge de sept ans, Wagner estime que la relation demeure jusqu'à l'âge de dix-huit ans:

Our data would indicate that the suggested cut-off point, viz. one or more, can be used to predict aggressive behavior once a subject has reached the age of eighteen. It will apparently be necessary to establish different cut-off points for ages below eighteen¹³.

Pour le contrôle du facteur de l'intelligence, Wagner conclut par une approche indirecte qu'il ne doit pas y avoir de corrélation significative entre le comportement agressif et le niveau d'intelligence. Ce point n'a pas encore été soumis à une étude systématique, aussi semble-t-il nécessaire d'accepter les résultats de son analyse a posteriori:

12 M. O. Oswald et P. Loftus, A Normative and Comparative Study of the Hand Test with Normal and Delinquent Children, dans Journal of Projective Techniques and Personality Assessment, vol. 31, no 4, livraison de sept. 1967, p. 62-68.

13 E. E. Wagner, B. Bricklin et Z. A. Piotrowski, op. cit., p. 86.

We feel an approximate degree of control has been exercised over the role of intelligence as it relates to Hand Test prediction for the following reasons:

- (1) The acting out score successfully differentiated the acting out inmates and the non acting out indigents who are comparable in normal or low normal intelligence;
- (2) The acting out score successfully differentiated the hospital acting out cases from the non acting out normals who are comparable in above normal intelligence;
- (3) The normals and indigents differ in intelligence, yet both groups do not act out, and both were predicted to be non acting out groups. The latter three points would indicate that the acting out score successfully predicts aggressive overt behavior over a wide range of intelligence levels, and that there is probably no great correlation between intelligence and the accuracy of the acting out scores¹⁴.

Bien que Drummond ait exclus de son échantillon les patients traités par chocs électriques et qu'il n'ait pas trouvé de variation notable pour la durée d'hospitalisation entre son groupe agressif et son groupe non agressif, Medvedeff a apparemment négligé de contrôler ces variables. L'effet de ces variables sur les tendances à l'agression est incertain.

Parmi les autres facteurs qui n'ont pas subi un contrôle rigoureux chez l'un et l'autre auteurs, on note l'évaluation du degré de schizophrénie et des traitements apportés aux patients (médicaments, thérapies, etc.). Une revue récente donne une importance particulière au contrôle de ces variables dans les recherches sur la schizophrénie.

¹⁴ Idem, ibidem, p. 85-86.

One of the most difficult task posed in research involving the mentally ill is that of making accurate measurements of the patient's overt behavior - his degree of psychopathology, his change as a result of treatment, his symptoms patterns. To meet this challenge, some investigators are directing their efforts toward the development of new rating scales, others toward the refinement of existing ones¹⁵.

Dû aux particularités diagnostiques de chaque hôpital et de chaque pays, il se pourrait qu'une population soit davantage schizophrénique que l'autre. Mais cette critique ne suffit pas à motiver le rejet des deux études car une approche clinique devrait montrer plus de flexibilité qu'une approche actuariale dans l'évaluation des indices de pathologie relatifs au contrôle des tendances agressives. Quant à contrôler la nature et l'intensité des traitements, on irait contre l'éthique professionnelle en privant le malade de ce qui pourrait améliorer son sort. De plus, un problème très réel se pose quand l'échantillonnage n'inclut que des patients non traités parce qu'ils ne semblent pas profiter des soins apportés :

¹⁵ National Institute of Mental Health, Department of Health, Education, and Welfare, U.S.A., Research in Schizophrenia, Public Health Service Publication No. 1175, livraison d'avril 1964, p. 13.

Investigators who compare groups of schizophrenics with normals should take into account the drug status of their patients. In addition, if they use patients who are not currently receiving drugs, they should attempt to obtain a sample which is not biased by selective exclusion of subjects from the sample. Any conclusions that are reached using only patients whose symptoms do not worsen with the withdrawal of tranquilizing drugs may be limited in their generality¹⁶.

C'est dire qu'un nouvel échantillonnage saurait difficilement obvier à ces difficultés.

Enfin, les deux auteurs ont employé le chi deux comme test de signification dans l'analyse statistique des résultats. Ce qui paraît adéquat puisque les échantillons se composent de catégories mutuellement exclusives. On remarque cependant l'absence de compte rendu de la constance interne et externe des correcteurs. Une nouvelle correction actuelle des protocoles remédiera à cette omission regrettable.

Bien que les études de Drummond et de Medvedeff ne répondent pas parfaitement aux normes de rigueur scientifique qui ont été exposées, on doit reconnaître qu'une troisième étude de contrevalidation saurait difficilement triompher des difficultés inhérentes à l'expérience. On ne voit donc pas d'objection à ce que la population de schizophrènes de la présente étude soit formée à partir de leurs données brutes.

¹⁶ L. I. Chapman, The Problem of Selecting Drug-Free Schizophrenics for Research, dans Journal of Consulting Psychology, vol. 27, n° 6, livraison de déc. 1963, p. 542.

CHAPITRE IV

LE PLAN D'ETUDE

Le chapitre précédent avait pour but de choisir le matériel projectif qui servira à la comparaison des méthodes actuaires et clinique. Il fut décidé que les études de Drummond et de Medvedeff allaient être à l'origine du matériel projectif.

On rencontre maintenant le problème de la sélection des juges. Deux principes très différents l'un de l'autre vont présider à la représentation du mode de prédiction clinique.

Le premier principe concerne la sélection de juges qualifiés et expérimentés afin de donner tous les avantages possibles à la méthode clinique dans sa recherche d'indices pertinents omis par le système actuaires présentement en usage.

Le second principe a trait à l'accessibilité de ces indices. Si l'on obtient les résultats attendus, une recherche subséquente visant à quantifier ces indices gagnerait à connaître au départ leur degré d'accessibilité.

Pour atteindre ce double objectif, on passera d'abord en revue les écrits d'importance qui traitent de la représentation de la méthode clinique. L'analyse des résultats obtenus dans ce domaine permettra peut-être de découvrir les règles d'une bonne représentation. Cette analyse conduira à

l'évaluation des difficultés inhérentes à la prédiction clinique avec le Hand Test. A la lumière des conclusions découlant de cet exposé, on proposera un mode de représentation de la méthode clinique dans la présente étude.

1. Représentations de la méthode clinique dans les écrits.

Une recherche faite par Cooke en 1967 porte sur les mérites de l'approche clinique pour la dérivation de formules actuaires. L'auteur souligne la nécessité d'employer un groupe de cliniciens expérimentés plutôt qu'un seul.

A comparison was made between the efficiency of using one skilled clinician versus a group of several skilled clinicians as a basis for deriving actuarial formulae. Both the reliabilities and validities of the pooled decisions of a group of clinicians were consistently higher than any single judge. Results suggest that any single judge who might have the highest validity on one occasion might not have the highest validity on another occasion¹.

Holt² avait déjà obtenu des résultats analogues en 1958.

La méthode clinique sera donc représentée par un groupe de juges dans le présent travail. Poussant plus loin

1 J. K. Cooke, Clinicians' Decisions as a Basis for Deriving Actuarial Formulae, dans Journal of Clinical Psychology, vol. 23, n° 2, livraison d'avril 1967, p. 232-233.

2 H. Holt, Clinical and Statistical Prediction: A Reformulation and Some New Data, dans Journal of Abnormal and Social Psychology, vol. 56, n° 1, livraison de janvier 1958, p. 1-12.

notre investigation, on peut se demander quels auraient été les résultats si Cooke avait employé des juges non expérimentés. Une revue sélective des écrits traitant du problème de la capacité prédictive en fonction du degré d'expérience des juges fait ressortir une tendance marquée.

Il semble tout d'abord que l'habileté à prédire soit une fonction hautement individuelle:

The studies here reviewed dealing with the process of diagnosis, stress the finding that among clinicians with roughly equivalent training, some seem able to perform at an easily demonstrable higher level of efficiency than others similarly trained and experienced³.

Il semble de plus que ce caractère spécifique de l'habileté à prédire se maintienne indépendamment du degré de sophistication des juges. L'étude de Walker, entr'autres, illustre bien cette tendance. Utilisant des juges à quatre niveaux de sophistication, des étudiants sous-gradués en génie, des étudiants sous-gradués en psychologie, des étudiants gradués en psychologie, et des psychologues expérimentés au niveau du doctorat, il arriva à ces conclusions:

³ J. Q. Holsopple et J. G. Phelan, The Skills of Clinicians in Analysis of Projective Tests, dans Journal of Clinical Psychology, vol. 10, n° 4, livraison d'oct. 1954, p. 307.

Conclusions drawn from an analysis of the data were (a) SCTs (sentence completion tests) were evaluated with consistency by even inexperienced judges though experienced judges did slightly better in terms of exact agreements in ratings, (b) all judges were able to classify SCT protocols as normal, neurotic or psychotic with an accuracy better than chance would allow, and (c) marked individual differences in judging ability appeared among judges of all levels of sophistication. Experienced judges were not necessarily more consistent or accurate, so it would appear that SCT data can be utilized by judges of limited psychological sophistication⁴.

Turner était déjà arrivé à des conclusions semblables en 1966. Utilisant quatre groupes de juges en psychologie (membres honoraires, jeunes docteurs, gradués et sous-gradués), il conclut que "no significant improvement in predictive accuracy occurred as the level of training and/or experience increased⁵."

Enfin, Johnston en 1967 corrobora une fois de plus ces résultats en demandant à des juges d'expériences et de professions différentes de prédire la durée d'hospitalisation chez des patients neuropsychiatriques: "neither professional background nor years of professional experience was related

⁴ C. E. Walker et J. D. Linden, Varying Degrees of Psychological Sophistication in the Interpretation of Sentence Completion Test Data, dans Journal of Clinical Psychology, vol. 23, n° 2, livraison d'avril 1967, p. 231.

⁵ D. R. Turner, Predictive Efficiency as a Function of Amount of Information and Level of Professional Experience, dans Journal of Projective Techniques and Personality Assessment, vol. 30, n° 1, livraison de février 1966, p. 5.

to predictive accuracy⁶."

Des résultats comparables ont été obtenus par les études indépendantes de Bialick⁷ en 1954, Grebstein⁸ en 1963 et Noddler⁹ en 1959. Weiss, pour sa part, découvrit en 1963 que des hommes de science en physique prédisent significativement moins bien que des docteurs en psychologie quand seulement un minimum d'informations descriptives est fourni, mais que leur niveau de fonctionnement dépasse celui atteint par les docteurs en psychologie quand le montant d'information fourni est sensiblement augmenté. Il résume ainsi les tendances qu'il a observées lors de sa revue des écrits:

6 R. Johnston et B. McNeal, Statistical versus Clinical Prediction: Length of Neuropsychiatric Hospital Stay, dans Journal of Abnormal Psychology, vol. 72, n° 4, livraison d'août 1967, p. 335.

7 I. Bialick et R. Hamlin, The Clinician as Judge: Details of Procedure in Judging Projective Material, dans Journal of Consulting Psychology, vol. 18, n° 4, livraison d'août 1954, p. 239-242.

8 L. Grebstein, Relative Accuracy of Actuarial Prediction, Experienced Clinicians, and Graduate Students in a Clinical Judgement Task, dans Journal of Consulting Psychology, vol. 27, n° 2, livraison d'avril 1963, p. 127-132.

9 E. B. Noddler et al., Objective Scoring versus Clinical Evaluation of the Bender-Gestalt, dans Journal of Clinical Psychology, vol. 15, n° 1, livraison de janvier 1959, p. 39-41.

Typically when the predictive performance of graduate students in psychology is compared with the performance of PhD psychologists, no significant differences in accuracy are found (Cline, 1955; Grigg, 1958; Kessen, 1957; Soskin, 1954, 1959). Only Klehr (1949) was able to report that experienced clinicians fared better than graduate students in predictive accuracy¹⁰.

Puisque le degré d'expérience, de sophistication et même la nature de la profession ne semblent pas influencer outre mesure l'habileté à prédire, le problème du choix des critères relatifs à la représentation de la méthode clinique dans le présent travail n'en est que plus aigu. On se voit alors dans l'obligation de recourir à une source d'information différente.

2. Exigences de la prédiction clinique avec le Hand Test.

On sait que Z. A. Piotrowski a participé au développement de la théorie et du système de correction du Hand Test. Selon lui, l'efficacité prédictive du clinicien sur un test projectif est liée de près à l'expérience pratique qu'il a de ce test.

¹⁰ J. H. Weiss, Effect of Professional Training and Amount of Information on Behavioral Prediction, dans Journal of Consulting Psychology, vol. 27, n° 3, livraison de juin 1963, p. 257.

Even when various projective tests give the same information, the rules of extracting that information from raw test data differ from test to test because each test has unique features, resulting from their unique test stimuli. Consequently, the degree of ease and validity with which information regarding certain specific behavior (especially overt behavior) is inferred does, and must differ from one personality test to another¹¹.

Un autre problème se pose, celui des qualifications requises des juges, indépendamment de leur familiarité avec le test. Au dire de Wagner,

Despite its relative simplicity, optimum efficiency in interpretation requires psychological sophistication. Its use should therefore be restricted to individuals at or beyond the graduate level who are conversant with personality dynamics and projective theory¹².

C'est dire que l'approche clinique serait avantageusement représentée par des juges au niveau gradué en psychologie et qui sont familiers avec le Hand Test. On peut enfin se demander si les juges dans leur travail de prédiction devront obéir à un schème mental particulier ou, encore, adopter une théorie spécifique de la personnalité. Toujours selon Wagner, il n'est pas nécessaire de contrôler ce facteur:

¹¹ Z. A. Piotrowski, The Rorschach Test, dans Prediction of Overt Behavior through the Use of Projective Techniques, Springfield, Illinois, Charles C. Thomas, 1960, p. 56.

¹² Edwin E. Wagner, The Hand Test, Manual for Administration, Scoring, and Interpretation, Akron, Ohio, Mark James, 1962, p. 2.

As with other projective techniques, the temptation to interpret responses in accordance with one's own theoretical predilections is both natural and alluring. There is nothing inherently reprehensible in integrating the Hand Test with a particular personality theory provided one avoids stretching the data out of shape on a pre-ordained, procrustean rack. Reasonable care should therefore be exercised to avoid doing extreme violence to the basic assumptions and empirical findings upon which this test is based¹³.

Cette citation de Wagner avait déjà été endossée par Piotrowski et Bricklin:

Any psychological theory and derived measure validly applicable to projective verbal production may be applied to Hand Test data¹⁴.

3. Présente représentation de la méthode clinique.

Sur le plan pratique, on rencontre une difficulté d'importance: bien que le Hand Test ait fait l'objet d'un nombre assez considérable d'études, il ne connaît qu'un usage restreint dans les milieux cliniques.

¹³ Idem, ibidem, p. 2.

¹⁴ E. E. Wagner, B. Bricklin et Z. A. Piotrowski, The Hand Test, A New Projective Test with Special Reference to the Prediction of Overt Aggressive Behavior, Springfield, Illinois, Charles C. Thomas, 1962, p. 94.

...new projective techniques may appear first in the literature and be research material before they have any general acceptance as clinically useful instruments. This is what appears to have happened to the Blacky (and possibly to the Bender-Gestalt used as a projective) and may be what is beginning to happen to the Hand Test and to the Holtzman. Again, research usage may anticipate later clinical usage¹⁵.

On peut alors rencontrer beaucoup de difficulté à obtenir dans un délai de temps raisonnable la collaboration d'un nombre suffisant de cliniciens expérimentés avec le test. Comme les écrits revus, Piotrowski et Wagner mis à part, accordent peu de valeur à l'expérience des juges avec un test, on pourrait s'en tenir à la sélection de juges non expérimentés avec le Hand Test.

D'un autre côté, on risque de mal représenter la méthode clinique en outrepassant des facteurs de sens commun, tel que le degré de sophistication et d'expérience des juges. Sans contredire les résultats obtenus à date, on peut se permettre de remarquer que les études revues ne se sont pas préoccupées des relations possibles entre les niveaux d'expérience et de sophistication d'une part, et des degrés de difficulté des tâches imposées aux juges d'autre part.

Pour ces raisons, on représentera la méthode clinique de manière à 1) respecter les règles posées par Wagner; 2)

¹⁵ D. H. Mills, The Research Use of Projective Techniques: a 17 Year Survey, dans Journal of Projective Techniques and Personality Assessment, vol. 29, n° 4, livraison de déc. 1965, p. 514-515.

tenir en ligne de compte la rareté des juges expérimentés avec le test; 3) accorder une certaine valeur à ce jugement de sens commun selon lequel le degré d'expérience et de sophistication influe sur l'efficacité prédictive du clinicien.

Voici comment la méthode clinique sera représentée pour les fins de la présente expérience. On formera trois groupes de juges, chaque groupe comptant dix juges. Les juges du premier groupe devront répondre aux critères d'acceptation suivant: expérience dans l'administration et l'interprétation du test; expérience avec les techniques projectives les plus connues, tels le Rorschach et le T.A.T.; grade académique minimum d'étudiant gradué en psychologie. Les juges du deuxième groupe auront à répondre aux mêmes exigences, avec cette différence qu'ils n'auront pas à être familiers avec le Hand Test. Enfin, le dernier groupe se composera de personnes ne possédant aucun diplôme en psychologie, aucune connaissance du test et des techniques projectives en général.

Le test t sera ensuite employé pour vérifier la présence ou l'absence de différences significatives entre ces groupes. Si les différences significatives trouvées sont en faveur du premier groupe, il sera nécessaire de rejeter les résultats du ou des groupes moins expérimentés. On aura alors à accroître le nombre de juges expérimentés avec le test.

Il semble que cette procédure puisse représenter une économie de temps et d'efforts sans que soient compromis les résultats de l'approche clinique. Elle peut en outre fournir des indications précieuses quant à l'accessibilité des indices recherchés.

CHAPITRE V

LE PLAN D'ETUDE

Le choix du matériel projectif qui va servir à la comparaison des approches actuariale et clinique a été fait; on a aussi suggéré un mode de représentation de la méthode clinique. Il reste tout d'abord à traiter de la sélection des protocoles des études de Drummond et de Medvedeff et de leur adaptation pour les fins de la présente expérience. On décrira ensuite les modalités de la tâche imposée aux cliniciens. Les définitions opératoires, hypothèses expérimentales et opérations statistiques seront présentées en dernier lieu.

1. Un échantillon stratifié de protocoles.

Pour que la tâche imposée aux cliniciens ne soit pas longue au point de provoquer un manque de motivation ou des effets indus de fatigue, il semble acceptable de choisir par stratification un nombre réduit de protocoles dans chaque étude: ainsi, des 136 protocoles originaux, soixante-et-quatre furent retenus de sorte que les études sont représentées par près de la demie de leur échantillon original. On sait de plus qu'au-delà de trente, les seuils de signification pour les valeurs de t et du χ^2 varient peu.

Le tableau I à la page suivante présente une comparaison entre l'échantillon obtenu par stratification et les échantillons originaux en termes de proportions, quant à la distribution des sexes et des nombres dans les sous-groupes agressifs et non agressifs. On constate que la moyenne des différences résultant de la stratification est de .000158 et que la plus grande variation, de l'ordre de .000383, se trouve dans la représentation actuelle du groupe non agressif de Medvedeff. Le tableau II permet une comparaison semblable entre la proportion de cas correctement prédits dans chaque sous-groupe dans le présent échantillon et dans les échantillons des deux études par le prédicteur d'agression et le score de non interaction, respectivement. La moyenne des différences résultant de la stratification est de .000360 et la plus grande différence observée, encore pour le groupe non agressif de Medvedeff, est de .000571.

2. Reproduction et adaptation des protocoles.

Les soixante-et-quatre protocoles sélectionnés furent dactylographiés à nouveau suivant une méthodologie nouvelle et uniforme afin d'enlever aux juges toute possibilité de discerner leur provenance. Les réponses verbales complètes furent reproduites. Les remarques faites par l'administrateur et les descriptions du comportement des sujets furent exclues afin de ne pas fournir aux juges des données

Tableau I.

Représentation des groupes agressifs (A) et non agressifs (NA) des études de Medvedeff et de Drummond dans l'échantillon actuel quant à la distribution des nombres et des sexes, en termes de proportions.

Echantillons Sexes	Medvedeff			Drummond		
	A	NA	Total	A	NA	Total
Echantillons originaux	.5000	.5000	1.0000	.5454	.4545	1.0000
Hommes	.2998	.2570	.5568	.2878	.2282	.5160
Femmes	.1999	.2427	.4426	.2575	.2282	.4857
Echantillon actuel	.5312	.4687	1.0000	.5625	.4374	1.0000
Hommes	.3125	.2187	.5312	.2812	.2187	.5000
Femmes	.2187	.2500	.4687	.2812	.2187	.5000

Tableau II.

Représentation des groupes agressifs (A) et non agressifs (NA) des études de Medvedeff et de Drummond dans l'échantillon actuel quant aux succès respectifs obtenus par le prédiseur d'agression (A.O.S.) et l'index de non interaction (WITH), en termes de proportions.

Echantillons Prédiseurs	Medvedeff			Drummond		
	A	NA	Total	A	NA	Total
Echantillons originaux						
<u>A.O.S.</u>	.5999	.7428	.6711	.4720	.5333	.5000
<u>WITH</u>	.7428	.7142	.7282	.3887	.5333	.4545
Echantillon actuel						
<u>A.O.S.</u>	.5882	.7999	.6897	.4444	.5000	.4687
<u>WITH</u>	.7646	.6666	.7187	.4444	.5000	.4687

particulières à la situation de testing.

Un numéro de code fut appliqué à chaque protocole pour en cacher l'identité, tant au point de vue de sa provenance que de sa nature agressive ou non agressive.

Les informations externes au test étaient limitées à l'âge et au sexe des patients. Les informations provenant du test même comprenaient les réponses verbales complètes, le ou les scores attachés à chaque réponse, les temps de réaction et la totalité des rapports numériques obtenus. Les deux raisons suivantes ont milité en faveur de l'inclusion du scoring complet: les cliniciens expérimentés avec le test auraient pu être influencés par leurs connaissances du système de correction et il aurait été dès lors impossible d'évaluer l'effet de cette variable sur les résultats obtenus; l'objectif du présent travail étant d'améliorer et non de supplanter le système actuel existant, il semble préférable de soumettre les juges à un examen constant des écarts qui existent entre les données actuelles et les informations recélées dans le matériel verbal.

L'évaluation des études de Drummond et de Medvedeff a fait ressortir l'absence de tout compte rendu de la constance interne et externe des correcteurs. Voilà pourquoi trois des juges expérimentés avaient pour tâche de recorriger chaque protocole au complet, sans l'aide de la correction originale, avant de procéder au travail d'évaluation clinique.

3. Modalités de la tâche imposée aux juges.

L'ordre de présentation des protocoles variait d'un juge à l'autre afin de décourager toute communication entre eux et également pour prévenir des effets de séquences possibles. Les variables de durée de la tâche, d'apprentissage gradué et de fatigue furent contrôlées en permettant aux juges de conserver les protocoles jusqu'à ce qu'ils soient satisfaits de leurs décisions.

Chaque juge avait pour tâche de classifier indépendamment les soixante-et-quatre protocoles comme agressifs ou non agressifs. Suivant la position prise par Isabelle, les juges furent forcés de choisir une alternative:

Prediction, using only extreme scores, may not be putting the clinician to the test. Furthermore, a significant comparison of the actuarial and clinical approaches to prediction, requires that prediction be made for all cases, or at least, on identical groups, by the two methods¹.

On eut recours à la procédure suivante afin de permettre l'évaluation de la constance interne. Pour chaque juge, le testing fut fait en deux temps, avec un minimum de trois jours d'intervalle. Un juge évaluait quarante-deux protocoles chaque fois. C'est dire que la seconde administration incluait vingt protocoles qui avaient déjà été

¹ L. Isabelle, Actuarial versus Clinical Methods in Predicting Achievement in a Science Faculty, thèse de doctorat présentée à l'École de Psychologie et d'Education, Ottawa, 1961, p. 62.

évalués. Le choix et l'ordre de présentation de ces vingt protocoles variaient d'un juge à l'autre suivant l'ordre du hasard.

Les juges avaient en main des instructions imprimées les instruisant sur la nature de la recherche en cours et sur les modalités du travail qu'ils avaient à accomplir. Le questionnaire d'information qui a servi à la division des juges en trois groupes y est aussi inclus. Le texte complet, trop long pour être reproduit ici, apparaît en appendice à la fin du travail.

Il devint bientôt évident que la question traitant de l'expérience du juge avec le Hand Test était par trop imprécise. On demanda alors aux juges de spécifier le nombre de Hand Test administrés et interprétés au moment de leur acceptation dans un groupe.

Faisant suite aux instructions, un modèle de feuille réponse illustre clairement la tâche à remplir. Ce modèle, reproduit en appendice, incitait les juges à donner par écrit les raisons majeures de chacune de leurs décisions. Une feuille réponse de ce genre suivait chaque protocole. On a dit que cette procédure pourrait permettre à une recherche subséquente de revenir sur ces décisions et de les quantifier si possible. Johnston et McNeal ont de plus observé une différence qualitative entre les explications données par leurs meilleurs juges et celles fournies par les juges de moindre

valeur, relativement à leurs décisions².

Enfin, une reproduction des neufs dessins du test, avec les réponses populaires correspondant à chaque stimulus, suivait la feuille réponse modèle. Cette feuille de location improvisée, que l'on retrouve à l'appendice 3, contient aussi des références au manuel du test à propos de l'analyse qualitative, de l'évaluation clinique de personnalités agressives et de schizophrènes. Ces précautions furent surtout prises à l'intention des juges moins expérimentés.

4. Définitions opératoires et hypothèses nulles.

On a vu que dans la méthode actuariale le prédisseur d'agression corrigé compare les tendances sociales positives de la personnalité, représentées par la somme des scores de dépendance, de communication et d'affection, aux tendances sociales négatives, représentées par la somme des scores de direction et d'agression. Un poids d'une unité est attaché à chaque score obtenu. La somme algébrique des tendances positives et des tendances négatives doit dépasser zéro pour que soit indiqué un excès de tendances agressives chez une personne. Drummond et Medvedeff ont tous deux emprunté cette méthode de Wagner pour le calcul du prédisseur d'agression

² R. Johnston et E. McNeal, Statistical versus Clinical Prediction: Length of Neuropsychiatric Hospital Stay, dans Journal of Abnormal Psychology, vol. 72, n° 4, livraison d'août 1967, p. 335-340.

corrigé.

La mesure de non interaction représentée par le Withdrawal Score se compose de la somme des réponses bizarres (BIZ), des réponses descriptives (DES) et des refus de répondre (FAIL). Un total supérieur à zéro indiquerait une incapacité d'entrer en interaction avec le milieu physique ou psychologique, que ce soit de façon positive ou négative.

Ces poids d'une unité, omniprésents dans le système de correction, n'ont pas été trouvés par analyse différentielle. C'est ce qu'illustre l'étude de Wiener³ qui a obtenu par une analyse a posteriori dans une population normale des résultats légèrement plus prononcés en modifiant le poids attaché à chaque score.

Une autre alternative serait de combiner ensemble les scores respectifs du prédictor d'agression et de l'index de non interaction pour permettre une prédiction actuariale plus réussie. Medvedeff déconseille cette procédure:

³ M. H. Wiener, The Hand Test and Prediction of Overt Aggression, thèse de maîtrise présentée à l'École de Psychologie et d'Education, Ottawa, 1965, vii-46 p.

A weighted combination of both AOS and WITH would have produced a much higher X^2 and would have decreased greatly the number of false positives and false negatives; however the use of some differentially weighted combination of the scores would be a posteriori and should be delayed pending further cross validation⁴.

Mais comme il n'entre pas dans les cadres du présent travail de participer directement à l'amélioration du système de correction, on se contentera de représenter fidèlement l'approche actuariale suivant son stage de développement atteint à date.

Pour faciliter la comparaison statistique des approches actuariale et clinique, la quantification des jugements obtenus par la méthode clinique se fera selon le procédé employé par Drummond et Medvedeff. C'est dire qu'un jugement exact obtiendra un point et qu'un jugement inexact obtiendra un score de zéro.

Les définitions opératoires des méthodes clinique et actuariale seront formulées ainsi pour les fins de la présente étude; la méthode actuariale sera représentée par le succès total obtenu par le prédicteur d'agression et l'index de non interaction, respectivement, dans les deux échantillons stratifiés; la méthode clinique sera représentée par le succès moyen obtenu par les juges dans chaque échantillon. On peut

⁴ E. E. Wagner et E. Medvedeff, Differentiation of Aggressive Behavior of Institutionalized Schizophrenics with the Hand Test, dans Journal of Projective Techniques, vol. 27, n° 1, livraison de mars 1963, p. 113.

maintenant formuler les sept hypothèses sous leur forme négative de différence nulle, que le présent travail se propose de mettre à l'épreuve:

1. Il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par le groupe de juges, et l'efficacité prédictive du prédisseur d'agression suivant la correction actuariale originale, dans les protocoles tirés de l'étude de Medvedeff.

2. Il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par le groupe de juges, et l'efficacité prédictive de l'index de non interaction suivant la correction actuariale originale, dans les protocoles tirés de l'étude de Medvedeff.

3. Il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par le groupe de juges, et l'efficacité prédictive du prédisseur d'agression suivant la correction actuariale originale, dans les protocoles tirés de l'étude de Drummond.

4. Il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par le groupe de juges, et l'efficacité prédictive de l'index de non interaction suivant la correction actuariale originale, dans les protocoles tirés de l'étude de Drummond.

5. Il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par le groupe de juges, et l'efficacité prédictive des deux prédisseurs d'agression suivant les corrections originales, dans les protocoles tirés des études de Medvedeff et de Drummond.

6. Il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par le groupe de juges, et l'efficacité prédictive des deux index de non interaction suivant les corrections originales, dans les protocoles tirés des études de Medvedeff et de Drummond.

7. Il n'y a pas de différence significative entre les résultats obtenus par la méthode clinique, représentée par le groupe de juges, dans l'étude de Medvedeff et ceux obtenus dans l'étude de Drummond.

5. Opérations statistiques.

L'évaluation de la constance actuariale des trois juges expérimentés agissant comme correcteurs s'est faite en termes de pourcentages d'accord des juges entr'eux et avec les corrections originales pour rendre possible la comparaison des résultats obtenus aux résultats rapportés dans les écrits. On employa la même procédure pour mesurer l'accord entre les diverses corrections du prédisseur d'agression et de l'index de non interaction. Le test du chi deux servit à sonder la valeur significative des différences observées. La constance interne fut calculée en pourcentages d'accord entre les deux corrections des vingt protocoles qui constituaient, pour chaque correcteur, le test-retest. On examina enfin la validité des diverses corrections actuariales du prédisseur d'agression et de l'index de non interaction au moyen du chi deux.

L'évaluation de la constance externe d'un juge fut faite en fonction de l'accord entre sa décision et les décisions prises par une majorité de juges pour chaque cas prédit. Une valeur de chi deux fut ensuite obtenue et traduite en coefficient phi. On eut alors recours au test t pour estimer la valeur significative des différences de moyennes entre le groupe de juges expérimentés et les deux groupes moins expérimentés. Le même procédé servit à la mesure de la constance interne des juges.

On eut de nouveau recours au test t pour sonder la signification des différences entre l'efficacité prédictive propre à chacun des trois groupes de juges. Le rapport F permit d'estimer les différences de variabilités des trois groupes, afin d'assurer la légitimité de l'emploi du test t .

Enfin, le test t fut appliqué aux résultats de l'expérience pour mettre à l'épreuve les sept hypothèses nulles décrites dans la section précédente.

CHAPITRE VI

PRESENTATION DES RESULTATS

Au chapitre précédent, la conduite de l'expérience a été décrite, de même que les hypothèses de recherche, les définitions opératoires et le choix des opérations statistiques. On doit maintenant rendre compte des résultats obtenus.

La constance et la validité des prédictions actuaires et clinique seront d'abord étudiées. Cet examen fera ressortir un inconvénient imprévu: les variations assez marquées, bien que non significatives, entre les présentes corrections actuaires et la correction originale de Medvedeff d'une part, et entre l'efficacité prédictive du premier groupe de juges et celles des deux groupes moins expérimentés, d'autre part, conduisent à une représentation supérieure de l'approche actuaire et à une représentation moyenne de l'approche clinique. Pour remédier à cette inégalité tout en respectant le plan de travail tracé au chapitre précédent, on permettra une représentation moyenne et supérieure de chaque approche.

Les résultats des quatre comparaisons entre les approches actuaires et clinique, supérieures et moyennes, seront alors présentées selon l'ordre des six premières hypothèses nulles déjà décrites. On éprouvera, en dernier lieu, la signification de la différence entre les résultats

obtenus par la méthode clinique supérieure et moyenne, respectivement, dans l'étude de Medvedeff et ceux obtenus dans l'étude de Drummond.

1. La constance des corrections actuaires.

Dans le but de comparer les résultats des diverses corrections actuaires de la présente expérience à ceux rapportés dans les écrits, le degré de concordance entre ces corrections fut calculé en termes de pourcentages suivant la procédure adoptée par Wagner:

In computing percentage agreements between the three scorers, "agreement" was defined as perfect identity on any of the fifteen specific scoring categories for each response. The final percentage represented the total number of agreements divided by the total number of scored responses, multiplied by 100 to eliminate decimals¹.

Les pourcentages obtenus par les trois correcteurs entr'eux et avec la correction originale par Drummond et Medvedeff sont consignés dans le tableau III. La moyenne générale est de 70%. Ce résultat semble bas si l'on se rapporte aux 80%, 78%, et 83% obtenus par Wagner² et au 90.5%

¹ E. E. Wagner, The Hand Test, Manual for Administration, Scoring, and Interpretation, Akron, Ohio, Mark James, 1952, p. 18.

² Idem, ibidem, p. 18.

Tableau III.

Le pourcentage d'accord des présentes corrections actuaires entrées et avec les corrections actuaires de Drummond et de Medvedeff.

Correcteurs	Juge 1	Juge 11	Juge 111
Drummond et Medvedeff	78.54%	66.58%	68.42%
Juge 1		68.49%	68.50%
Juge 11			65.17%

de Foucar-Egyed³. Oswald n'avait cependant obtenu que 73% entre deux corrections indépendantes. Voici l'interprétation qu'il accorde à ce résultat.

While this may be regarded as satisfactory, one would expect on this basis to incorrectly score about 3 responses in a protocol of 12⁴.

La constance interne des trois correcteurs de notre étude fut de 92%, 77% et 73%, ce qui donne une moyenne de 81%.

Les résultats obtenus remettent en question la valeur de la constance externe des correcteurs. Pour cette raison, une évaluation fut faite de la constance du prédiseur d'agression et de l'index de non interaction, selon la procédure établie par Wagner. La moyenne de ces pourcentages, paraissant au tableau IV, atteint cette fois-ci 70% pour le prédiseur d'agression et 83% pour l'index de non interaction. Les moyennes correspondantes pour l'évaluation de la constance interne sont de 88% et de 87%. Le tableau V rapporte la constance interne de chaque correcteur sur le prédiseur d'agression et l'index de non interaction dans les deux études.

³ C. Foucar-Egyed, Overt Aggression Expressed between the Ages of 6-9-13, as Defined by Scores on The Hand Test, thèse de Maîtrise présentée à l'École de Psychologie et d'Éducation, Ottawa, 1966, p. 63.

⁴ M. O. Oswald et P. Loftus, A Normative and Comparative Study of the Hand Test with Normal and Delinquent Children, dans Journal of Projective Techniques and Personality Assessment, vol. 31, n° 4, livraison de septembre 1967, p. 68.

Tableau IV.

Le pourcentage d'accord entre les diverses corrections actuelles du prédiseur d'agression (A.O.S.) et de l'index de non interaction (WITH).

Correcteurs	A.O.S.			WITH		
	Juge I	Juge II	Juge III	Juge I	Juge II	Juge III
Drummond et Medvedeff	82.81%	68.75%	64.06%	84.38%	78.13%	85.94%
Juge I		75.00%	65.63%		79.69%	84.38%
Juge II			62.50%			87.50%

Tableau V.

La constance interne des correcteurs de la présente expérience sur le prédiseur d'agression (A.O.S.) et sur l'index de non interaction (WITH) en termes de pourcentages.

Correcteurs	Prédiseurs	
	<u>A.O.S.</u>	<u>WITH</u>
Juge I	100.%	95.%
Juge II	90.%	85.%
Juge III	75.%	80.%

On sonda ensuite la valeur significative des différences entre les correcteurs quant à leur maniement du prédisseur d'agression et de l'index de non interaction. La formule⁵ particulière suivante du chi deux fut employée pour mettre cette hypothèse à l'épreuve:

$$\chi^2 = \frac{(A-D)^2}{A+D}$$

La correction de Yates fut employée quand les fréquences théoriques étaient inférieures à dix.

Les valeurs de chi deux ($df=1$ et $N=64$) sont rapportées au tableau VI. On remarque la présence d'une différence significative à .02 entre la correction originale et les juges II et III, une différence significative à .01 entre les juges I et III, et une différence significative à .001 entre les juges II et III sur le prédisseur d'agression.

Les valeurs de chi deux obtenues pour l'index de non interaction ne sont pas significatives au niveau de .05. C'est dire que l'index de non interaction demeure relativement stable d'un correcteur à l'autre, alors que le prédisseur d'agression varie de façon significative dans presque tous les cas.

⁵ L.-T. Dayhaw, Manuel de Statistique, Ottawa, Editions de l'Université d'Ottawa, 1963, p. 393.

Tableau VI.

Les valeurs de chi deux correspondant à l'accord des diverses corrections actuaires du prédiseur d'agression (A.C.S.) et de l'index de non interaction (WITH).

Correcteurs	A.C.S.			WITH		
	Juge I	Juge II	Juge III	Juge I	Juge II	Juge III
Drummond et Medvedeff	.400	6.400*	6.545*	.444	.285	.000
Juge I		2.250	8.909**		.077	.400
Juge II			16.666***			.000

* Une valeur de chi deux significative à .02.

** Une valeur de chi deux significative à .01.

*** Une valeur de chi deux significative à .001.

2. La validité des corrections actuaires.

Le test de signification du chi deux fut appliqué aux résultats de chaque correction actuariale afin de sonder la valeur significative de sa prédiction. Ces valeurs de chi deux paraissant au tableau VII doivent être interprétées en fonction de l'hypothèse nulle suivante: les corrections actuaires, prises une à une, ne permettent pas de prédire au delà de la chance les cas agressifs et non agressifs au moyen, respectivement, du prédicteur d'agression et de l'index de non interaction. La formule employée fut celle du chi deux dans un tableau à quatre cases calculé directement des fréquences de cases en tenant compte de la correction de Yates pour continuité.

On note que parmi les valeurs de chi deux obtenues, seule celle de la correction originale de Medvedeff pour l'index de non interaction s'avère significative à .05 ($df=1$ et $N=32$). La valeur correspondante qu'il rapporte dans son étude était significative à .001 avec un chi deux de 12.87 ($df=1$ et $N=70$). Une différence semblable se trouve pour le prédicteur d'agression entre le chi deux original de 7.06 ($df=1$ et $N=70$), significatif à .01, dans l'étude de Medvedeff, et le chi deux de 3.50 ($df=1$ et $N=32$), non significatif à .05, résultant des trente-deux protocoles tirés de son échantillon. Comme le pourcentage de cas correctement prédits par le prédicteur d'agression et l'index de non interaction demeure

Tableau VII.

La signification des diverses corrections actuaires du pré-
dicteur d'agression (A.O.S.) et de l'index de non interaction
(WITH) dans les protocoles tirés des études de Drummond et
de Medvedeff, en valeurs de chi deux.

Correcteurs	Etude de Medvedeff		Etude de Drummond	
	<u>A.O.S.</u>	<u>WITH</u>	<u>A.O.S.</u>	<u>WITH</u>
Drummond			.002	.002
Medvedeff	3.500	4.400 [*]		
Juge I	.1104	.003	.019	.019
Juge II	3.071	.002	.018	.019
Juge III	.4481	.086	.004	.019
Wagner			.013	.073

* Une valeur de chi deux significative à .05.

essentiellement le même de l'échantillon original au présent échantillon, ces différences doivent résulter de la grandeur de N qui est passé de soixante-dix à trente-deux. Il faut donc interpréter les valeurs du tableau VII avec précaution. Si l'on porte le N de trente-deux à soixante-dix, en augmentant proportionnellement les valeurs de cases, on obtient alors un chi deux de 8.13, significatif à .01, pour le prédicteur d'agression correspondant au juge II, et un chi deux de 2.49, non significatif à .05, pour le prédicteur d'agression correspondant au juge III. Les autres valeurs obtenues par les correcteurs de la présente expérience étant inférieures à celle du juge III pour le prédicteur d'agression, et celle-ci n'étant pas significative à .05, on peut conclure que le fait d'élever N à soixante-dix apporte deux différences significatives en faveur de Medvedeff et une différence significative en faveur du juge II pour le prédicteur d'agression; toutes les autres valeurs demeurent non significatives au niveau de .05.

Dans le but de pousser plus loin l'analyse des divergences entre les résultats des diverses corrections actuelles, la formule précédente du test du chi deux fut employée pour éprouver l'hypothèse nulle suivante: la tendance à prédire correctement au moyen du prédicteur d'agression et de l'index de non interaction, respectivement, ne varie pas significativement d'un correcteur à l'autre. Le N étant de soixante-quatre (4x4) et aucune fréquence théorique n'étant inférieure

à cinq, la correction de Yates ne fut pas appliquée.

Les résultats consignés aux tableaux VIII et IX ne font ressortir aucune différence significative au niveau de .05. On constate tout de même que les prédictions faites au moyen de la correction de Medvedeff furent plus réussies. Quant aux quatre prédictions actuaires dans l'étude de Drummond, les valeurs de chi deux sont minimales et n'atteignent pas le niveau de signification .5. C'est dire que la méthode actuariale sera représentée à son meilleur si les corrections originales servent à la comparaison des méthodes actuariale et clinique.

3. La constance de la méthode clinique.

L'évaluation de la constance externe d'un juge fut faite en fonction de l'accord entre sa décision et les décisions de la majorité des juges pour chaque cas prédit, sans égard à la validité de ces décisions. On trouvera le pourcentage ainsi obtenu par chaque juge dans l'appendice rapportant les caractéristiques individuelles des juges. La moyenne de ces vingt-six pourcentages est de 72.59%. Une formule⁶ du chi deux sert à éprouver la valeur significative de cette moyenne:

$$\chi^2 = \frac{2(f_o - f_t)^2}{f_t}$$

⁶ L.-T. Dayhaw, Op. cit., p. 379.

Tableau VIII.

La signification des différences d'efficacité prédictive entre les diverses corrections actuaires du prédiscur d'agression (A.O.S.) et de l'index de non interaction (WITH) des protocoles tirés de l'étude de Medvedeff, en valeurs de chi deux*.

Correcteurs	A.O.S.			WITH		
	Juge I	Juge II	Juge III	Juge I	Juge II	Juge III
Medvedeff	1.067	0.000	1.067	1.696	2.400	3.217
Juge I		1.067	0.000		0.063	0.247
Juge II			1.067			0.626

* Une valeur de chi deux de 3.84 serait requise pour une différence significative à .05 de probabilité.

Tableau IX.

La signification des différences d'efficacité prédictive entre les diverses corrections actuelles du prédiscur d'agression (A.C.S.) et de l'index de non interaction (WITH) des protocoles tirés de l'étude de Drummond, en valeurs de chi deux*.

Correcteurs	A.O.S.			WITH		
	Juge I	Juge II	Juge III	Juge I	Juge II	Juge III
Drummond	0.000	0.000	0.000	0.254	0.254	0.254
Juge I		0.000	0.000		0.000	0.000
Juge II			0.000			0.000

* Une valeur de chi deux de 3.84 serait requise pour une différence significative à .05 de probabilité.

La valeur de 12.25 ($df=1$ et $N=26$) ainsi obtenue s'avère significative au niveau de .001. Utilisant cette formule⁷ de conversion,

$$\chi^2 = N\phi^2$$

un coefficient phi de .44 est trouvé. Le test t fut ensuite employé pour évaluer la signification des différences de moyennes entre le groupe de juges expérimentés et les deux groupes non expérimentés.

La valeur de t ($df=5$ et $N=6$) résultant de la comparaison du premier groupe et du deuxième groupe, dont les moyennes respectives sont de 46.33 et de 48.20, en scores bruts, fut de .5436 et non significative au niveau de .05.

La valeur correspondante ($df=5$ et $N=6$) pour la comparaison du premier groupe et du troisième groupe, dont la moyenne était de 44.8, fut de .445, et non significative à .05.

Le même procédé fut employé pour l'évaluation de la constance interne des juges. Le pourcentage obtenu par chaque juge paraît à l'appendice 4. La moyenne de ces pourcentages est de 74% et correspond à un coefficient phi de .48, significatif à .05. L'évaluation de la différence entre la moyenne de 15.33, en scores bruts, du premier groupe, et la moyenne de 14.90 du deuxième groupe rendit un t de 1.14 ($df=5$ et $N=6$), non significatif à .05. La seconde comparaison, entre la

7 L.-T. Dayhaw, op. cit., p. 401.

moyenne du premier groupe et celle du troisième groupe ($M = 14,40$) donna une valeur de t de 2,46 ($dl=5$ et $N=6$), non significative à .05.

4. L'efficacité prédictive des trois groupes de juges comparés.

Pour éprouver l'hypothèse d'une différence significative entre l'efficacité prédictive du groupe de juges expérimentés et celles des deux groupes de juges moins expérimentés, pris un à un, on eut recours à la formule pour le test t d'une différence entre deux moyennes non en corrélation.

La comparaison de la moyenne de 38,66, en score brut, du premier groupe à celle de 35,40 du deuxième groupe donna un t de 1,88 ($dl=14$ et $N_1+N_2=6+10$); et la comparaison du premier groupe au troisième groupe ($M = 35,00$), un t de 2,12 ($dl=14$ et $N_1+N_2=6+10$). Ces résultats ne sont pas significatifs au niveau de .05, qui requiert une valeur de t de 2,145.

Comme l'emploi du test t suppose des variances à peu près égales d'un groupe à l'autre pour le cas d'une différence entre deux statistiques, le rapport F sert à évaluer la valeur significative des différences entre les variabilités des groupes comparés.

Les F ($dl=9$ et 5) de 0,939 et de 0,935 obtenus respectivement pour la première et deuxième comparaison, ne sont pas significatifs au niveau de .05.

Puisqu'aucune différence significative ne fut obtenue entre la constance ou l'efficacité prédictive du premier groupe et la constance ou l'efficacité prédictive des deux groupes moins expérimentés, pris un à un, on ne voit pas d'objection à ce que les trois groupes de juges soient fondus en un seul groupe homogène, pour représenter l'approche clinique dans la comparaison statistique des approches actuariale et clinique.

On a déjà remarqué que si les corrections actuaires originales, par Drummond et Medvedeff, sont employées pour la comparaison des méthodes actuariale et clinique, la méthode actuariale est alors représentée à son meilleur. Or le fait de fondre les trois groupes de juges en un seul conduit à une représentation moyenne de la méthode clinique. La méthode clinique semble donc désavantagée. Les considérations qui suivent ont pour but de remédier à cette situation.

Bien que les différences entre les trois groupes de juges ne sont pas significatives, il se pourrait que l'habileté à prédire de façon clinique soit une fonction hautement individuelle et relativement indépendante de l'expérience avec le test ou des connaissances en psychologie. La grande majorité des écrits revus au chapitre IV, qui traitait des diverses représentations de la méthode clinique, supporte cette hypothèse. On peut ainsi représenter la méthode clinique supérieure dans cette étude par les juges qui ont fait

preuve d'un rendement nettement supérieur, sans tenir compte de leur groupe d'appartenance.

Le test t et le rapport F servirent à nouveau pour éprouver l'hypothèse d'une différence significative entre l'efficacité prédictive du groupe de juges supérieurs, représenté par les neuf meilleurs juges, et celles des trois groupes de juges.

Les valeurs de t résultant de la comparaison entre la moyenne du groupe supérieur et les moyennes du deuxième et troisième groupe furent de 3.57 ($dl=17$ et $N_1+N_2=9+10$) et de 3.97 ($dl=17$ et $N_1+N_2=9+10$) et significatives, respectivement, à .01 et .001. La valeur de t provenant de la comparaison du groupe de juges supérieurs au premier groupe n'étant que de 1.02 ($dl=13$ et $N_1+N_2=9+6$), elle ne fut pas significative à .05. Enfin, on ne trouva pas de différence significative au niveau de .05 pour les F de .286 ($dl= 9$ et 8), de .287 ($dl= 9$ et 8) et de .305 ($dl= 5$ et 8) obtenus respectivement dans la première, deuxième et troisième comparaison.

Cette procédure conduit donc à une double représentation de la méthode clinique et permettra de comparer: l'approche actuariale supérieure et l'approche clinique moyenne; l'approche actuariale supérieure et l'approche clinique supérieure; l'approche actuariale moyenne et l'approche clinique moyenne; l'approche actuariale moyenne et l'approche clinique supérieure. Ce nombre plus grand de comparaisons permettra sans doute une

interprétation plus différenciée des résultats.

5. Comparaison des méthodes actuaires supérieure et clinique moyenne.

La première hypothèse à éprouver, sous sa forme négative de différence nulle, se formulait ainsi: il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par les vingt-six juges, et l'efficacité prédictive du prédiseur d'agression suivant la correction actuariale originale dans les protocoles tirés de l'étude de Medvedeff. Utilisant la formule

$$t = \frac{M_1 - M_2}{\sqrt{\frac{\sum(X - M)^2}{N(N-1)}}}$$

un t ($df=25$ et $N=26$) de 19.04 fut obtenu en faveur du prédiseur d'agression. Un t de 3.725 était requis pour que la différence soit significative au niveau de .001.

Substituant l'index de non interaction au prédiseur d'agression dans l'analyse qui précède, on obtient un t ($df=25$ et $N=26$) de 27.74 en faveur de la méthode actuariale. Ce résultat est significatif à .001.

On eut recours aux mêmes opérations statistiques pour éprouver la troisième et la quatrième hypothèse nulle: il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique moyenne, et l'efficacité prédictive du prédiseur d'agression et de l'index de non interaction,

respectivement, suivant la correction actuariale originale, dans les protocoles tirés de l'étude de Drummond.

Un t de 1.94 ($df=25$ et $N=26$), non significatif à .05, fut obtenu en faveur de la méthode clinique dans la première comparaison. Un résultat exactement semblable fut obtenu pour la seconde comparaison.

Les cinquième et sixième hypothèses ont été formulées ainsi, sous leur forme négative de différence nulle: il n'y a pas de différence significative entre l'efficacité prédictive de la méthode clinique, représentée par les vingt-six juges, et l'efficacité prédictive du prédicteur d'agression et de l'index de non interaction, respectivement, suivant la correction originale, dans les protocoles tirés des études de Drummond et de Medvedeff.

On obtint cette fois-ci, encore au moyen de la formule employée pour éprouver les quatre premières hypothèses, un t de 1.428 ($df=25$ et $N=26$), non significatif à .05 de probabilité, en faveur du prédicteur d'agression, et un t de 2.857 ($df=25$ et $N=26$), significatif à .01, en faveur de l'index de non interaction. On trouvera au tableau X une comparaison des diverses moyennes obtenues, ainsi que la signification des différences entre les moyennes des prédictions cliniques et des prédictions actuaires.

Tableau X.

Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiseur d'agression (A.O.S.) et l'index de non interaction (WITH) dans les corrections actuaires originales et par le groupe total de juges, et la signification des différences en valeurs de "t".

Etudes	Moyennes "t"			Moyennes "t"		
	Juges	<u>A.O.S.</u>		Juges	<u>WITH</u>	
Medvedeff	19.81	22.00	19.04**	19.81	23.00	27.74**
Drummond	16.19	15.00	1.94	16.19	15.00	1.94
Medvedeff et Drummond	36.00	37.00	1.43	36.00	38.00	2.86*

* Une valeur de "t" significative au niveau de .01 de probabilité.

** Une valeur de "t" significative au niveau de .001 de probabilité.

6. Comparaison des méthodes actuaires et cliniques supérieures.

Les juges représentant la méthode clinique supérieure étaient au nombre de neuf. La moyenne des prédictions réussies par ce groupe était de 62.50%. Le pourcentage maximum de 65.63% fut rejoint par trois juges; aussi, trois juges étaient à la limite inférieure de 59.06%.

Remplaçant les résultats de l'approche clinique moyenne par ceux de l'approche clinique supérieure, on mit à l'épreuve les six hypothèses mentionnées dans la section précédente. Le tableau XI présente une comparaison des moyennes et des valeurs de t ($df=8$ et $N=9$) résultant des différences entre les moyennes actuaires et cliniques.

Le tableau fait ressortir la présence de six différences significatives, dont quatre au compte de l'approche clinique. Il semble bien que les deux différences significatives au niveau de .001 obtenues dans l'étude de Medvedeff par la méthode actuaires sont compensées par les deux différences significatives, également au niveau de .001, remportées par la méthode clinique dans l'étude de Drummond. De plus, la prédiction clinique dans les études combinées dépasse les prédictions actuaires correspondantes et s'avère significative au niveau de .02.

Tableau XI.

Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiscur d'agression (A.O.S.) et l'index de non interaction (WITH) dans les corrections actuaires originales et par le groupe de juges supérieurs, et la signification des différences en valeurs de "t".

Etudes	Moyennes "t"			Moyennes "t"		
	Juges	A.O.S.		Juges	WITH	
Medvedeff	20.33	22.00	6.33***	20.33	23.00	10.11***
Drummond	19.67	15.00	6.23***	19.67	15.00	6.23***
Medvedeff et Drummond	40.00	37.00	4.99**	40.00	38.00	3.33*

- * Une valeur de "t" significative à .02.
- ** Une valeur de "t" significative à .01.
- *** Une valeur de "t" significative à .001.

7. Comparaison des méthodes actuariale et clinique moyennes.

La comparaison des résultats obtenus par les approches actuariale et clinique moyennes dans l'étude de Medvedeff, dans l'étude de Drummond, et dans les deux études combinées, permit cette fois-ci de déceler trois différences significatives ($df=25$ et $N=26$) en faveur de l'approche clinique.

La première différence, significative au niveau de .01, ressort de la comparaison des moyennes obtenues par la méthode clinique et l'index de non interaction dans l'étude de Drummond. La deuxième différence, significative à .001, découle de la comparaison des moyennes obtenues par l'approche clinique et l'index de non interaction dans l'étude de Medvedeff. La troisième différence, significative à .001, fut obtenue par la comparaison des deux mêmes modes de prédiction dans les études de Drummond et de Medvedeff combinées.

Les six valeurs de t correspondant à la signification des différences de moyennes sont rapportées au tableau XII. On constate que l'approche actuariale n'obtint qu'une différence minime en sa faveur, non significative à .05, au moyen du prédicteur d'agression dans l'étude de Medvedeff.

8. Comparaison des méthodes actuariale moyenne et clinique supérieure.

Si l'on représente l'approche clinique à son mieux, les résultats obtenus dépassent ceux atteints par le prédicteur

Tableau XII.

Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiseur d'agression (A.O.S.) et l'index de non interaction (WITH) dans les diverses corrections actuaires et par le groupe total de juges, et la signification des différences en valeurs de "t".

Etudes	Moyennes "t"			Moyennes "t"		
	Juges	A.O.S.		Juges	WITH	
Medvedeff	19.81	20.00	1.65	19.81	18.50	11.39**
Drummond	16.19	15.20	1.61	16.19	14.00	3.57*
Medvedeff et Drummond	36.00	35.00	1.43	36.00	32.00	5.71**

* Une valeur de "t" significative à .01.
 ** Une valeur de "t" significative à .001.

d'agression et l'index de non interaction dans l'approche actuariale moyenne. Les résultats et la signification des différences provenant des diverses comparaisons sont consignées au tableau XIII. On voit que toutes les valeurs de t ($df=8$ et $N=9$) furent significatives au niveau de .001, sauf celle de 1,25, non significative à .05, obtenue dans la comparaison du prédisseur d'agression au mode de prédiction clinique dans l'étude de Medvedeff.

Les tendances observées à date portent à croire que le mode de prédiction clinique conserve son efficacité indépendamment du degré de difficulté de la tâche. Cette affirmation pour l'instant gratuite mérite une tentative de validation.

9. La prédiction clinique dans les études de Drummond et de Medvedeff.

La méthode clinique étant d'abord représentée par le groupe total de juges, elle obtint un score brut de 515 dans l'étude de Medvedeff et un score brut de 421 dans l'étude de Drummond. La signification des différences entre les deux séries de mesure fut éprouvée au moyen du test t . On eut recours à cette formule particulière⁸ qui tient compte de la corrélation pouvant exister entre les deux séries de mesure.

⁸ L.-T. Dayhaw, op. cit., p. 364.

Tableau XIII.

Comparaison des moyennes obtenues dans Drummond et Medvedeff, respectivement et ensemble, par le prédiscur d'agression (A.O.S.) et l'index de non interaction (WITH) dans les diverses corrections actuelles et par le groupe de juges supérieurs, et la signification des différences en valeurs de "t".

Etudes	Moyennes "t"			Moyennes "t"		
	Juges	A.O.S.		Juges	WITH	
Medvedeff	20.33	20.00	1.25	20.33	18.50	6.93*
Drummond	19.67	15.20	5.96*	19.67	14.00	7.56*
Medvedeff et Drummond	40.00	35.00	8.32*	40.00	32.00	13.31*

* Une valeur de "t" significative à .001.

$$t = \frac{M_D}{\frac{(D - M_D)^2}{N(N-1)}}$$

Un t ($dl=25$ et $N=26$) de 3.68 fut obtenu. La différence entre les paires de mesures s'avéra donc significative au niveau de .01 de probabilité.

La méthode clinique étant ensuite représentée par le groupe de juges supérieurs, des scores bruts de 183 et de 177 furent respectivement obtenus dans les études de Medvedeff et de Drummond. Comme le t ($dl=8$ et $N=9$) de 0.456 n'atteint pas la valeur requise à .05 de probabilité, on ne put rejeter l'hypothèse nulle voulant qu'il n'y ait pas de différence significative entre les deux séries de mesures pour le groupe de juges supérieurs.

Bien qu'on ne puisse en toute rigueur prouver la non existence d'une relation significative, les résultats mentionnés suggèrent que les juges supérieurs maintiennent un niveau d'efficacité uniforme indépendamment des difficultés de la tâche. Il n'en est pas de même pour le groupe total de juges.

CHAPITRE VII

L'ANALYSE DES RESULTATS

On a présenté au chapitre précédent les résultats des diverses comparaisons entre les méthodes actuaires et cliniques, supérieures et moyennes. Il s'agit à présent d'analyser ces résultats pour en faire ressortir les tendances générales. La constance et la validité prédictive du prédicteur d'agression, de l'index de non interaction et de l'approche clinique, seront étudiées tour à tour. On procédera ensuite à l'évaluation des mérites relatifs à chaque mode de prédiction dans les diverses comparaisons des méthodes actuaires et cliniques. Enfin, les conclusions générales qui s'imposent seront tirées et interprétées à la lumière des écrits revus aux chapitres précédents.

1. L'évaluation de la constance.

Une étude de la constance obtenue dans les divers modes de prédiction montre tout d'abord que la constance interne de la méthode actuaires, avec une moyenne de 81%, dépassait la constance interne de la méthode clinique de 6%. D'autre part, la constance externe des juges étant de 73%, elle fut légèrement supérieure à la constance externe du prédicteur d'agression (70%) mais nettement inférieure à celle de l'index de non interaction (83%).

Eprouvant la signification des différences entre correcteurs d'après leur maniement du prédiseur d'agression et de l'index de non interaction, respectivement, on a vu que les cinq différences significatives obtenues dévalorisaient la constance du prédiseur d'agression. Quant aux résultats en faveur de l'index de non interaction, ils sont d'une portée restreinte puisque l'emploi de ce prédiseur semble limité à des populations de schizophrènes. C'est dire que le clinicien moyen profiterait de respecter davantage la prédiction actuare au moyen de l'index de non interaction lorsqu'il s'agit d'évaluer les tendances à l'agression de schizophrènes. Ses prédictions seront ainsi plus constantes; mais seront-elles mieux réussies?

2. L'évaluation de la validité.

Bien que les différences d'efficacité prédictive entre correcteurs ne furent pas significatives, un seul correcteur réussit à atteindre, au moyen du prédiseur d'agression, le seuil de prédiction obtenu par la correction originale de Medvedeff. On peut se demander si les résultats rapportés par Medvedeff avec le prédiseur d'agression et l'index de non interaction surestiment l'efficacité prédictive du test, ou si les résultats des nouvelles corrections ne sont pas plutôt des sousévaluations. En admettant que ces deux possibilités soient partiellement vraies, la représentation

moyenne du mode de prédiction actuariaire devient alors très acceptable.

Les différences d'efficacité prédictive entre les trois groupes de juges ne furent pas significatives. Le groupe de juges supérieurs prédit cependant à un niveau significativement plus élevé que les deuxième et troisième groupes. Comme la différence n'était pas significative entre le groupe de juges supérieurs et le groupe de juges expérimentés, on peut supposer qu'avec une meilleure représentation du groupe de juges expérimentés, les différences entre ce groupe et les deuxième et troisième groupes auraient pu être significatives.

De toute façon, un juge sur trois prédit significativement mieux que le juge moyen. Sur les trois nouvelles corrections actuaires du prédicteur d'agression et de l'index de non interaction, dans l'étude de Medvedeff, une seule correction du prédicteur d'agression fut significative à .01 avec une valeur de chi deux moindre que celle obtenue par l'auteur. Dans l'étude de Brummond, les présentes corrections n'approchent même pas le seuil d'une différence significative à .05. Il paraît ainsi plus difficile en général de prédire avec succès au moyen d'une approche actuariaire qu'au moyen d'une approche clinique: l'approche clinique rejoint une fois sur trois le seuil maximal apparent de validité prédictive, tandis que l'approche actuariaire n'y parvient qu'une fois sur six.

3. Les comparaisons des modes de prédiction.

Pour simplifier l'évaluation des mérites propres à chaque approche, on étudiera d'abord les résultats des comparaisons entre l'approche clinique et le prédiseur d'agression supérieurs et moyens. Les résultats des comparaisons entre l'approche clinique et l'index de non interaction supérieurs et moyens seront ensuite examinés. Il s'agira enfin d'évaluer globalement les mérites respectifs des approches actuariale et clinique. Afin de ne pas encombrer le lecteur, les niveaux de signification ne seront pas rapportés: l'expression "différence significative" comprendra les niveaux de .01 et de .001 de probabilité.

Un total de douze tests de signification furent calculés entre les résultats respectifs du prédiseur d'agression et de l'approche clinique supérieurs et moyens. On ne trouva que deux différences significatives en faveur du prédiseur d'agression mais quatre différences significatives au compte de la prédiction clinique.

Les deux modes de prédiction moyens semblent être d'une égale efficacité. Cependant, le prédiseur d'agression supérieur dépasse significativement l'approche clinique moyenne dans l'étude de Medvedeff; par contre, la prédiction clinique moyenne dépasse significativement les résultats du prédiseur d'agression moyen dans l'étude de Drummond, et dans les deux études combinées. Enfin, trois différences

significatives ressortent des comparaisons entre les approches actuaires et clinique supérieures: la première, en faveur du prédisseur d'agression dans l'étude de Medvedeff; les deux autres, en faveur des juges supérieurs dans l'étude de Drummond et dans les études de Medvedeff et de Drummond combinées.

Au moins trois tendances ressortent de ces comparaisons: les approches actuaires et clinique moyennes sont d'une égale efficacité; l'approche clinique supérieure sait mieux profiter de son avantage lorsqu'on l'évalue en fonction du prédisseur d'agression moyen; les prédisseurs d'agression moyen et supérieur sont significativement moins valides que les prédictions cliniques supérieures quand la tâche inclut les protocoles des deux études.

Les différences d'efficacité prédictive entre l'approche clinique supérieure et moyenne et l'index de non interaction supérieur et moyen, au nombre de douze, furent significatives à huit occasions en faveur de l'approche clinique et à trois occasions en faveur de l'index de non interaction.

On note cette fois un succès complet pour les approches cliniques moyenne et supérieure comparées au prédisseur actuaires moyen. C'est dire que les trois succès actuaires proviennent de l'index de non interaction supérieur: la correction originale de Medvedeff lui permet ainsi de dépasser significativement l'approche clinique moyenne dans l'étude de Medvedeff et dans les études de Drummond et de Medvedeff considérées

ensemble, et l'approche clinique supérieure dans l'étude de Medvedeff. Par contre, la prédiction clinique supérieure, comparée à l'index de non interaction supérieur, remporte deux différences significatives, la première dans l'étude de Drummond, l'autre dans les deux études combinées.

Examinant la portée de ces résultats, on conclut que l'approche clinique démontre une applicabilité plus générale que l'index de non interaction. Au plan pratique, le clinicien moyen qui ne peut donner qu'une correction actuariale moyenne de ce prédicteur y gagnerait à utiliser son jugement clinique; le clinicien supérieur, capable de tirer le maximum de l'index de non interaction en se limitant à une correction actuariale, y gagnerait encore à employer son jugement deux fois sur trois. Et si la correction actuariale de ce clinicien atteint à peine ou ne dépasse guère la moyenne des corrections actuariales, il ferait bien de s'en tenir exclusivement à une appréciation clinique des protocoles.

En résumé, sur les vingt-quatre valeurs de t obtenues dans les diverses comparaisons, dix-sept furent significatives: cinq étaient au compte de la prédiction actuariale supérieure; l'approche clinique supérieure en obtint neuf; et l'approche clinique moyenne en obtint trois. Quant à la prédiction actuariale moyenne, elle se montre d'une efficacité égale ou, plus souvent, d'une efficacité inférieure aux approches cliniques moyenne et supérieure.

4. La portée des résultats.

Un fait s'impose à prime abord, celui de la complexité des problèmes inhérents à la quantification d'un instrument projectif tel que le Hand Test. Borgotta, qui a influencé l'orientation du présent travail, insiste sur l'importance de la quantification:

Quantification is essential if there is to be an increasing probability of accurate prediction based upon fewer and more precisely defined sets of variables and if the same predictions are to be made by all trained people and not merely by the sensitive few¹.

Mais les résultats de la présente recherche indiquent au contraire qu'il est plus aisé de trouver un bon clinicien qu'un bon correcteur dans une tâche de prédiction du comportement agressif de schizophrènes au moyen du Hand Test. Pour résoudre ce paradoxe apparent, il semble à propos de se rappeler la pensée de Fisher, qui croit que des résultats contradictoires dans les tentatives de validation d'un test projectif résultent souvent du degré précaire de développement atteint par le système de correction actuair en usage.

Puisque les approches cliniques supérieure et moyenne ont dépassé de beaucoup les approches actuaires supérieure

¹ E. P. Borgatta et H. Philip, The Definition of Some Problem Areas for Research: a Theoretical Formulation of some Problems of Relevance to Diagnostics, dans Group Psychotherapy, vol. 6, n° 1 et 2, livraison de mai et août 1953, p. 101.

et moyenne dans la différenciation des schizophrènes agressifs et non agressifs, on conclut qu'un montant substantiel d'information contenu dans le matériel projectif des protocoles employés dans cette étude n'a pas été incorporé dans le système de correction du test. Les raisons écrites par les juges supérieurs pour chacune de leurs décisions réussies pourraient peut-être permettre à une recherche subséquente de quantifier ces informations. Il serait alors intéressant de répéter l'expérience en représentant le mode de prédiction actuariale par le système de correction ainsi amélioré.

Deux considérations s'imposent pour faire le lien entre les résultats de la présente étude et l'assertion catégorique de Meehl qui dénigre le mode de prédiction clinique et accorde tous les avantages au mode de prédiction actuariale, quelle que soit sa valeur: aussi objective qu'elle paraisse, la correction actuariale d'un test projectif demeure un instrument limité qui doit passer par différents stades de développement; de plus, les adeptes de l'une et l'autre approches devraient sans doute prendre conscience des avantages respectifs de chaque mode de prédiction au lieu de s'attarder à des victoires de parti pris.

En dernier lieu, quelques recommandations à l'intention des chercheurs qui voudront pousser plus loin le travail de validation du Hand Test: pour permettre un contrôle plus grand des variables expérimentales, on pourrait recourir à

une mesure quantitative du degré de pathologie dans la population étudiée; si l'on tente une contrevalidation du présent travail, il paraît nécessaire tout d'abord d'employer une mesure externe à la tâche pour la sélection du groupe de juges supérieurs et, enfin, d'exiger un nombre comparable d'évaluations actuelles et d'évaluations cliniques.

RESUME ET CONCLUSIONS

Le but principal de cette étude était de comparer les mérites relatifs des approches actuariale et clinique sur le Hand Test dans une tâche de prédiction du comportement agressif.

Le matériel projectif qui a servi à la comparaison des méthodes actuariale et clinique consistait en un échantillon stratifié de protocoles de schizophrènes tirés des études de Drummond et de Medvedeff. Les vingt-six juges représentant le mode de prédiction clinique furent d'abord répartis en trois groupes, d'après leurs connaissances avec les techniques projectives en général et leur expérience avec le Hand Test en particulier; un dernier groupe fut enfin formé, qui comprenait les juges ayant fait preuve d'une habileté prédictive supérieure. Le Withdrawal Score et Acting Out Score étaient les prédicteurs d'agression actuaires. Les quatre corrections des protocoles rendirent possible une représentation moyenne et une représentation supérieure du mode de prédiction actuariale. On procéda alors à une comparaison statistique des méthodes actuariale et clinique supérieures et moyennes.

Bien que l'expérience avec le Hand Test et avec les tests projectifs les plus connus n'ont pas influencé de façon significative l'efficacité prédictive des juges, les résultats des diverses comparaisons favorisent, en général, le

mode de prédiction clinique; sur un total de dix-sept différences significatives au niveau de .01 de probabilité, douze différences étaient à l'avantage de la méthode clinique, dont neuf significatives au niveau de .001. Une autre conclusion se dégageant de ces résultats serait que les juges supérieurs maintiennent une efficacité relativement stable, en dépit des variations en difficulté d'une tâche à l'autre. On note cependant un gain significatif au compte des prédicteurs actuels supérieurs dans l'étude de Medvedeff.

La présente étude ne prétend pas prouver la supériorité incontestable du mode de prédiction clinique sur le mode de prédiction actuariaire. D'une part, les résultats de l'expérience ne sont pas tous concluants; d'autre part, les raisons écrites qui ont motivé les décisions réussies par les juges supérieurs pourraient permettre à une recherche subséquente d'améliorer le test comme instrument de prédiction de l'agression. Comme toute comparaison des mérites relatifs des approches actuariaire et clinique est intimement liée au stage de développement atteint par l'approche actuariaire, les mérites relatifs des deux approches comme moyens de prédiction du comportement agressif sur le Hand Test et dans une population de schizophrènes, seraient à nouveau matière à recherche.

On devrait alors obtenir une mesure quantitative du degré de pathologie dans la population étudiée afin de

permettre à d'autres recherches un contrôle plus grand des variables expérimentales. De préférence, on devrait aussi employer une mesure externe à la tâche pour la sélection des juges supérieurs et exiger un nombre comparable d'évaluations actuelles et d'évaluations cliniques.

BIBLIOGRAPHIE

Borgatta, E. F. et H. Philip, The Definition of some Problem Areas for Research: a Theoretical formulation of some problems of relevance to diagnostics, dans Group Psychotherapy, vol. 6, n° 1 et 2, livraison de mai et août 1953, p. 90-101.

Un écrit qui a influencé l'orientation de base du présent travail: "Quantification is essential if there is to be an increasing probability of accurate predictions..."

Cooke, J. K., Clinicians' Decisions as a Basis for Deriving Actuarial Formulas, dans Journal of Clinical Psychology, vol. 23, n° 2, livraison d'avril 1967, p. 232-233.

L'auteur étudie les mérites de l'approche clinique dans la dérivation de formules actuaires. Une conclusion de l'expérience souligne la nécessité d'employer un groupe de cliniciens expérimentés plutôt qu'un seul.

Drummond, F., A Failure in the Discrimination of Aggressive Behavior of Undifferentiated Schizophrenics with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 30, n° 3, livraison de juin 1966, p. 276-279.

L'auteur n'a pas réussi à différencier des schizophrènes agressifs de schizophrènes non agressifs en dépit des différences significatives en faveur du WITH et du A.O.S. rapportées par Medvedeff et Wagner en 1963. Il ne voit pas d'explication valable à cette divergence de résultats.

Fisher, S., Projective Methodologies, dans Annual Review of Psychology, vol. 18, 1967, p. 165-185.

Les tentatives de validation d'un test dont le système de correction est en voie de développement peuvent souvent conduire à des résultats apparemment contradictoires. Voilà une des conclusions que tire Fisher lors de sa critique de valeur sur le développement des techniques projectives.

Girardin, N. B., The Use of the Hand Test for Differentiating Simple and Paranoid Schizophrenics, thèse de maîtrise présentée à l'école de Psychologie et d'Education, Ottawa, 1964, viii-52 p.

Le système de correction du test, sauf pour les réponses bizarres, ne permet pas de différencier des schizophrènes simples et paranoïaques. L'auteur suggère une évaluation qualitative des protocoles pour déceler les différences entre les deux groupes.

Gleser, G. C., faisant la revue du Hand Test dans la section Character and Personality (p. 141-540) de The Sixth Mental Measurement Yearbook, O. K. Buros, Highland Park, The Gryphon Press, N.-J., 1965, p. 436-438.

Une appréciation critique du test qui met l'accent sur ses potentialités encore inexploitées.

Grigg, A. E., Experience of Clinicians, and Speech Characteristics, and Statements of Clients as Variables in Clinical Judgement, dans Journal of Consulting Psychology, vol. 22, n° 4, livraison d'août 1958, p. 315-319.

Un effort de recherche en vue de contrôler des variables souvent négligées dans les études sur le jugement clinique.

Helsopple, J. Q. et J. G. Phelan, The Skills of Clinicians in Analysis of Projective Tests, dans Journal of Clinical Psychology, vol. 10, n° 4, livraison d'octobre 1954, p. 307-320.

Une recension des écrits traitant de l'habileté à prédire de façon clinique et une étude menée par l'auteur soulignent le caractère individuel de cette habileté.

Holt, R., Clinical and Statistical Prediction: A Reformulation and Some New Data, dans Journal of Abnormal and Social Psychology, vol. 56, n° 1, livraison de janvier 1958, p. 1-12.

En réaction aux critiques de Meehl, l'auteur cherche à réévaluer la méthode clinique. Il clarifie certains termes et propose de nouvelles définitions opératoires.

Medvedeff, E. et E. E. Wagner, Differentiation of Aggressive Behavior of Institutionalized Schizophrenics with the Hand Test, dans Journal of Projective Techniques, vol. 27, n° 1, livraison de mars 1963, p. 111-113.

Des schizophrènes agressifs furent différenciés de schizophrènes non agressifs au moyen du WITH et du A.C.S. aux niveaux respectifs de .001 et de .01.

Meehl, P. E. When Shall We Use Our Head Instead of the Formula?, dans Journal of Counseling Psychology, vol. 4, n° 4, livraison d'hiver 1957, p. 268-273.

Suivant la publication de son livre, cette article contribue à séparer davantage les adeptes des approches actuaire et clinique. L'approche clinique est considérée comme étant un mode de prédiction de dernier recours.

National Institute of Mental Health, Department of Health, Education and Welfare, U.S.A., Research in Schizophrenia, Public Health Service Publication n° 1175, livraison d'avril 1964, y-18 p.

Une synthèse très documentée sur les diverses orientations prises par les chercheurs s'intéressant à l'étude de la schizophrénie.

Oswald, M. O. et P. Loftus, A Normative and Comparative Study of the Hand Test with Normal and Delinquent Children, dans Journal of Projective Techniques and Personality Assessment, vol. 31, n° 4, livraison de septembre 1967, p. 62-68.

Une tentative de validation chez des enfants normaux et délinquents des résultats positifs rapportés par Wagner et ses disciples. L'auteur croit que le A.O.S. doit être interprété en fonction du groupe social composant l'échantillon étudié.

Sawyer, J., Measurement and Prediction, Clinical and Statistical, dans The Psychological Bulletin, vol. 66, n° 3, 1966, p. 178-200.

Une analyse soignée de plusieurs travaux importants liés à la controverse entre les modes de prédiction actuaire et clinique. Les conclusions apportées facilitent la compréhension des divergences apparentes et réelles.

Shaw, D. L. et J. D. Linden, A Critique of the Hand Test, dans Educational and Psychological Measurement, vol. 24, n° 2, livraison d'été 1964, p. 283-284.

Une appréciation critique du test qui remet en question sa valeur prédictive.

Wagner, E. E., B. Bricklin et Z. A. Piotrowski, The Hand Test, A New Projective Test with Special Reference to the Prediction of Overt Aggressive Behavior, Springfield, Illinois, Charles C. Thomas, 1962, x-100 p.

La première publication majeure sur le test: une source d'information importante qui rapporte les résultats des premières tentatives de validation.

Wagner, E. E., The Hand Test, Manual for Administration, Scoring, and Interpretation, Akron, Ohio, Mark James, 1962, 62 p.

Suivant de près la monographie, le manuel présente un système de correction amélioré, des modèles de correction et d'interprétation.

Weiss, J. H., Effect of Professional Training and Amount of Information on Behavioral Prediction, dans Journal of Consulting Psychology, vol. 27, n° 3, livraison de juin 1963, p. 257-262.

On y trouve une recension incomplète des écrits traitant du problème de l'efficacité prédictive chez des cliniciens de divers niveaux d'expérience.

Watzel, H., R. J. Shapiro et E. R. Wagner, Prediction of Recidivism among Juvenile Delinquents with the Hand Test, dans Journal of Projective Techniques and Personality Assessment, vol. 31, n° 4, livraison de septembre 1967, p. 69-72.

Bien que les délinquents occasionnels et habitués soient différenciés au delà de la chance par le A.C.S. et le score d'agression, l'auteur suggère de considérer l'intensité de la psychopathologie et non pas seulement le seuil des tendances à l'agression.

APPENDICE 1
INSTRUCTIONS

INSTRUCTIONS

These protocols have been selected from two studies, both attempting to predict aggressive behaviour in undifferentiated schizophrenics with the Hand Test. It was hypothesized in both studies that an acting-out score (AOS) of +1 or greater indicates some aggressive proclivities, and that a withdrawal score (WITH) of +1 or greater is a contraindication of aggression. The subjects were assigned to the aggressive (A) or non aggressive (NA) group on the basis of their past history and behaviour on the ward. The criteria used for aggression are the following: hostile acts, destruction of property, attempts to escape, verbal abuse, refusal to take medication, refusal to work when able to. The criteria for non aggressive behaviour were as follows: general passivity, easy cooperation, willingness to work and absence of record of willful destruction. Doubtful cases were excluded. The results obtained from the American study are positive, showing a success frequency of 67% when using the AOR and of 71% when using the WITH score, in correctly classifying aggressive and non aggressive patients. In spite of even more stringent criteria for aggression in the English study, the corresponding figures were only 50% and 45%. There is no obvious reason to explain these significantly different results.

The present study aims to improve predictive accuracy by allowing judges to alter the classification made on the basis of the AOR and/or the WITH whenever they feel it is worthwhile on any type of evidence they might find useful, as long as each judge is solely responsible for his judgements. Each judge is therefore requested to classify each protocol as either A or NA. He may use any reference he wishes from the literature. In order to cancel any learning, fatigue and time effects, he is welcome to keep the protocols until he is thoroughly satisfied with his scorings, even if it implies going over the same protocols several times.

Closely linked to the ongoing controversy of actuarial versus clinical prediction, our rationale is that the Hand Test numerical ratios, the WITH and the AOR, are not sufficiently accurate predictors of acting out at this time to do without careful clinical judgement.

Each judge will be informed of his success frequency and the success frequency of the group upon request.

Please answer these questions as accurately as possible:

Name: _____ Age: _____ Phone number: _____

Academic standing:

-M.A. level: what year? _____ Post M.A. experience: _____ mths _____ yrs

-Ph.D. level: what year? _____ Post Ph.D. experience: _____ mths _____ yrs

Do you have any experience:

- in the interpretation of projective tests? Yes No
- in the interpretation of the Hand Test? Yes No
- with schizophrenics patients? Yes No
- are you in favor of clinical prediction? Yes No

N.B.: please do not write on the protocols, use the answer sheets.

Thanking you for your kind cooperation,

W. Rodrigue.

APPENDICE 2

SAMPLE ANSWEE SHEET

SAMPLE ANSWER SHEET

SEX: F

AGE: 40

CODE: XXX

Your DECISION:

AGGRESSIVE: XNON AGGRESSIVE:

MAIN REASON(S) MOTIVATING YOUR DECISION:

The response to card n reveals the presence of marked verbal hostility. However, because of the evaluative nature of the response, it could not be scored as such. The short reaction times (impulsivity?), the large number of FAIL responses (poor cooperation?) and the absence of clear AFF responses suggest that this woman may well be aggressive in spite of an AOS of -1 (i.e. an acting out ratio, AOR, of 3:2) and of a WITH score of 4. Perhaps in this case the WITH score, heavily weighted with BIZ responses, reflects the delusional direction of her aggressive tendencies rather than a true non aggressive orientation.

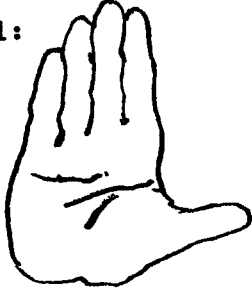
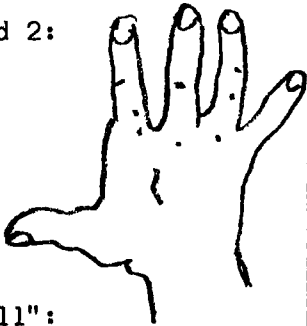

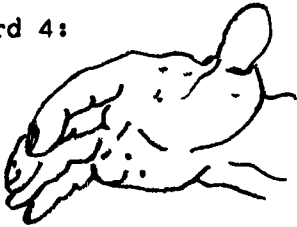





YOUR NAME: Judge JOHN SMITH

YOUR IDENTIFYING NUMBER: 100

APPENDICE 3

REPRODUCTION OF THE 9 HAND TEST CARDS

Reproduction of the 9 Hand Test cards (the tenth is blank) and the "popular responses" for each "card pull"¹

<p>Card 1:</p>  <p>"pull": DIR or AFF</p>	<p>Card 2:</p>  <p>"pull": ACQ, ACT, PAS</p>	<p>Card 3:</p>  <p>"pull": ACT, DIR</p>
<p>Card 4:</p>  <p>"pull": none</p>	<p>Card 5:</p>  <p>"pull": ACQ, ACT, PAS</p>	<p>Card 6:</p>  <p>"pull": AGG</p>
<p>Card 7:</p>  <p>"pull": none</p>	<p>Card 8:</p>  <p>"pull": ACT</p>	<p>Card 9:</p>  <p>"pull": ACT</p>
<p>¹ For complete information on qualitative interpretation, cf. Wagner's <u>Manual</u> (content analysis: p. 28-29; "card pull": p. 30-31; verb and adjective usage: p. 29; schizophrenia: p. 32-33; antisocial personalities: p. 34).</p>		

APPENDICE 4
CARACTERISTIQUES INDIVIDUELLES DES JUGES

Caractéristiques individuelles des juges.

A. Les six juges expérimentés^a avec le Hand Test et avec les tests projectifs les plus connus:

1. Age: 25
 Sexe: M
 Occupation: Psychomètre
 Niveau académique^b: M.A. (Ps), thèse incomplète
 Expérience du Hand Test: 75
 Constance interne: 90%
 Constance externe: 64%
 Succès obtenu: 62%
 En faveur de la prédiction clinique: oui

2. Age: 24
 Sexe: F
 Occupation: Orienteur
 Niveau académique: M.A. (Ps), thèse incomplète
 Expérience du Hand Test: 200
 Constance interne: 75%
 Constance externe: 73%
 Succès obtenus: 55%
 En faveur de la prédiction clinique: oui

3. Age: 42
 Sexe: M
 Occupation: Etudiant
 Niveau académique: Ph.D., première année
 Expérience du Hand Test: 10
 Constance interne: 70%
 Constance externe: 78%
 Succès obtenu: 55%
 En faveur de la prédiction clinique: oui

^a L'expérience du test fut mesurée en terme du nombre total approximatif d'administrations et d'interprétations.

^b Les termes M.A. (Ps), M.Ps., et Ph.D. correspondent aux définitions rapportées dans l'Annuaire de la Faculté de Psychologie et d'Education de l'Université d'Ottawa.

4. Age: 23
 Sexe: M
 Occupation: Etudiant
 Niveau académique: M.A. (Ps), troisième année
 Expérience du Hand Test: 10
 Constance interne: 70%
 Constance externe: 69%
 Succès obtenu: 63%
 En faveur de la prédiction clinique: "when the tool allows it"
5. Age: 29
 Sexe: M
 Occupation: Psychologue clinicien
 Niveau académique: Ph.D. complété en 1967
 Expérience du Hand Test: 30
 Constance interne: 80%
 Constance externe: 67%
 Succès obtenu: 63%
 En faveur de la prédiction clinique: oui
6. Age: 33
 Sexe: F
 Occupation: Psychologue clinicien
 Niveau académique: Ph.D. complété en 1965
 Expérience du Hand Test: 500
 Constance interne: 80%
 Constance externe: 83%
 Succès obtenu: 55%
 En faveur de la prédiction clinique: oui

B. Les dix juges expérimentés^b avec les tests projectifs les plus connus:

1. Age: 33
 Sexe: M
 Occupation: Psychologue clinicien
 Niveau académique: Ph.D. complété en 1967
 Expérience du Hand Test: aucune
 Constance interne: 90%
 Constance externe: 77%
 Succès obtenu: 46%
 En faveur de la prédiction clinique: oui

^b Un juge ne devait pas avoir plus de deux administrations du Hand Test pour appartenir à ce groupe.

2. Age: 27
 Sexe: M
 Occupation: Etudiant
 Niveau académique: M.Ps., troisième année
 Expérience du Hand Test: aucune
 Constance interne: 90%
 Constance externe: 80%
 Succès obtenu: 63%
 En faveur de la prédiction clinique: non
3. Age: 25
 Sexe: M
 Occupation: Etudiant
 Niveau académique: M.A. (Ps), troisième année
 Expérience du Hand Test: aucune
 Constance interne: 85%
 Constance externe: 88%
 Succès obtenu: 56%
 En faveur de la prédiction clinique: oui
4. Age: 26
 Sexe: M
 Occupation: Etudiant
 Niveau académique: Ph.D., première année
 Expérience du Hand Test: aucune
 Constance interne: 75%
 Constance externe: 73%
 Succès obtenu: 61%
 En faveur de la prédiction clinique: oui
5. Age: 23
 Sexe: M
 Occupation: Orienteur
 Niveau académique: M.Ps., complété en 1967
 Expérience du Hand Test: 1
 Constance interne: 90%
 Constance externe: 77%
 Succès obtenu: 51%
 En faveur de la prédiction clinique: oui
6. Age: 20
 Sexe: F
 Occupation: Etudiante
 Niveau académique: M.Ps., troisième année
 Expérience du Hand Test: aucune
 Constance interne: 75%
 Constance externe: 75%
 Succès obtenu: 56%
 En faveur de la prédiction clinique: oui

7. Age: 25
 Sexe: M
 Occupation: Etudiant
 Niveau académique: M.Ps., troisième année
 Expérience du Hand Test: aucune
 Constance interne: 55%
 Constance externe: 75%
 Succès obtenu: 56%
 En faveur de la prédiction clinique: oui
8. Age: 23
 Sexe: M
 Occupation: Etudiant
 Niveau académique: M.Ps., troisième année
 Expérience du Hand Test: aucune
 Constance interne: 40%
 Constance externe: 66%
 Succès obtenu: 55%
 En faveur de la prédiction clinique: oui
9. Age: 28
 Sexe: M
 Occupation: Orienteur
 Niveau académique: M.Ps., complété en 1965
 Expérience du Hand Test: 2
 Constance interne: 65%
 Constance externe: 69%
 Succès obtenu: 57%
 En faveur de la prédiction clinique: oui
10. Age: 24
 Sexe: M
 Occupation: Orienteur
 Niveau académique: M.Ps., complété en 1967
 Expérience du Hand Test: aucune
 Constance interne: 80%
 Constance externe: 75%
 Succès obtenu: 54%
 En faveur de la prédiction clinique: oui

C. Les dix juges ne possédant aucune connaissance des tests projectifs:

1. Age: 23
Sexe: F
Occupation: Etudiante en peinture
Niveau académique: 10^e année
Constance interne: 80%
Constance externe: 72%
Succès obtenu: 62%
En faveur de la prédiction clinique: oui
2. Age: 61
Sexe: F
Occupation: Service civil-archiviste
Niveau académique: 12^e année
Constance interne: 60%
Constance externe: 63%
Succès obtenu: 46%
En faveur de la prédiction clinique: oui
3. Age: 34
Sexe: F
Occupation: Infirmière
Niveau académique: Diplôme en sciences infirmières
Constance interne: 55%
Constance externe: 72%
Succès obtenu: 54%
En faveur de la prédiction clinique: oui
4. Age: 46
Sexe: F
Occupation: Secrétaire
Niveau académique: 12^e année
Constance interne: 80%
Constance externe: 66%
Succès obtenu: 54%
En faveur de la prédiction clinique: oui
5. Age: 28
Sexe: F
Occupation: Secrétaire
Niveau académique: 12^e année
Constance interne: 65%
Constance externe: 73%
Succès obtenu: 56%
En faveur de la prédiction clinique: oui

6. Age: 23
 Sexe: F
 Occupation: Secrétaire
 Niveau académique: 12^e année
 Constance interne: 90%
 Constance externe: 77%
 Succès obtenu: 51%
 En faveur de la prédiction clinique: oui
7. Age: 28
 Sexe: M
 Occupation: Professeur
 Niveau académique: Bac. en Commerce
 Constance interne: 70%
 Constance externe: 61%
 Succès obtenu: 55%
 En faveur de la prédiction clinique: oui
8. Age: 24
 Sexe: F
 Occupation: Thérapeute occupationnelle
 Niveau académique: B.A. et diplôme en thérapie occupationnelle
 Constance interne: 60%
 Constance externe: 61%
 Succès obtenu: 56%
 En faveur de la prédiction clinique: oui
9. Age: 41
 Sexe: F
 Occupation: Ménagère
 Niveau académique: 11^e année et cours d'enseignement ménager
 Constance interne: 95%
 Constance externe: 78%
 Succès obtenu: 49%
 En faveur de la prédiction clinique: oui
10. Age: 25
 Sexe: F
 Occupation: Institutrice
 Niveau académique: Brevet B, école normale
 Constance interne: 65%
 Constance externe: 78%
 Succès obtenu: 63%
 En faveur de la prédiction clinique: oui

APPENDICE 5

ABSTRACT OF

Prédictions actuaires et clinique du comportement
agressif chez des schizophrènes au moyen du Hand Test

ABSTRACT OF

Prédictions actuaires et clinique du comportement
agressif chez des schizophrènes au moyen du Hand Test¹

Meehl's hypothesis that actuarial prediction is more accurate than clinical prediction was studied by means of a comparison between actuarial versus clinical prediction on the Hand Test. Validation studies on this test heretofore have produced conflicting results, which have cast considerable doubt as to whether the test has achieved its primary goal, namely the prediction of overt aggressive behavior.

This paradox, according to Fisher, is likely a consequence of an inadequately developed scoring system.

The study was designed to evaluate the relative accuracy of an average and a high actuarial correction of the test predictors of aggression versus the accuracy of two groups of judges of varying clinical ability in differentiating a sample of hospitalized schizophrenic patients into "aggressive" and "non aggressive".

The results obtained generally favored the clinical method. It was concluded that a substantial amount of information could possibly be quantified in order to improve the actuarial predictions, and that, until then, a clinician

¹ William Rodrigue, Thèse de maîtrise présentée à la Faculté de Psychologie de l'Université d'Ottawa, 1968, x-123.

is well advised to rely on his subjective evaluations as well as statistical findings.

Certain limitations in the study were pointed out, and suggestions for further validation of the test were advanced.

APPENDICE 6

SOMMAIRE DE

Prédictions actuaire et clinique du comportement
agressif chez des schizophrènes au moyen du Hand Test

SOMMAIRE DE

Prédictions actuaire et clinique du comportement
agressif chez des schizophrènes au moyen du Hand Test¹

Les modes de prédiction actuaire et clinique furent comparés au moyen du Hand Test dans le but d'éprouver l'hypothèse de Meehl voulant que la méthode actuaire soit d'une efficacité prédictive supérieure à la méthode clinique. Comme les tentatives de contrevalidation de ce test ont obtenu des résultats contradictoires à maintes reprises, on doute qu'il ait atteint son premier objectif, celui de prédire le comportement agressif.

Ce paradoxe, selon Fisher, pourrait résulter d'un système de correction inadéquat parce qu'en voie développement.

Le plan d'expérience de la présente étude permettait de comparer l'efficacité des prédicteurs d'agression du test, suivant une correction actuaire moyenne et supérieure, à l'efficacité prédictive d'un groupe de juges supérieurs et d'un groupe de juges moyens, pour différencier un échantillon de schizophrènes hospitalisés en cas "agressifs" et "non agressifs".

¹ William Rodrigue, Thèse de Maîtrise présentée à la Faculté de Psychologie de l'Université d'Ottawa, 1968, x-123.

Les résultats obtenus favorisent en général le mode de prédiction clinique. On conclut alors qu'un clinicien profite d'évaluer subjectivement les prédisseurs d'agression actuelles du test à la lumière du protocole complet, du moins jusqu'à ce que le système de correction soit amélioré.

Après avoir souligné les limitations inhérentes à l'étude, certaines suggestions furent présentées à l'intention des chercheurs qui voudront pousser plus loin le travail de validation du test.