

UNE ETUDE CRITIQUE ET COMPARATIVE DU CANADIAN INTELLIGENCE EXAMINATION

par Eva Labrosse

Thèse présentée à la faculté des arts
de l'Université d'Ottawa par l'inter-
médiaire de l'Institut de Psychologie
en vue de l'obtention de la maîtrise
ès arts.

Ottawa, Canada, 1946

Université d'Ottawa
BIBLIOTHÈQUES



LIBRARIES
University of Ottawa



UMI Number: EC55379

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 EC55379
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 a été préparée sous la direction du directeur de l'Institut de Psychologie de l'Université d'Ottawa, le R.F. Raymond-Henri Shevenell, O.I.I.

Nous lui exprimons notre sincère gratitude pour son précieux concours.

TABLE DES MATIÈRES

Chapitre	page
INTRODUCTION	vi
I.- LE <u>CANADIAN INTELLIGENCE EXAMINATION</u> SELON AMOSS	1
1. La nécessité de ce test individuel	1
2. Le but que ce proposait l'auteur	4
3. La conception de l'intelligence selon Amoss	5
4. Fidélité à la révision <u>Stanford</u>	7
5. Les principales modifications apportées à la révision <u>Stanford-Binet</u>	9
II.- ETUDE CRITIQUE ET COMPARATIVE DE CHAQUE SÉRIE D'ITEMS.	21
1. Les dessins à copier	25
2. Le pliage de feuilles	35
3. Les images	40
4. La répétition de chiffres à la suite	49
5. La répétition de chiffres à reculons	57
6. La répétition de syllabes	60
7. Les définitions	62
8. Le test de vocabulaire	65
9. Les ressemblances et les différences	71
10. Les problèmes d'arithmétique	77
11. Les séries mathématiques	81
12. Compter des <u>sous</u>	83
13. Les jours de la semaine et les mois de l'année	85
14. Reproduire des dessins de mémoire	87
III.- ETUDE CRITIQUE ET COMPARATIVE DES ITEMS PARTICULIERS	93
1. Nommer des objets	94
2. Les formes géométriques	96
3. La compréhension	97
4. Nommer quatre couleurs	99
5. Le jeu de patience	102
6. Exécuter trois commissions	105
7. La droite et la gauche	107
8. Reconnaître des pièces de monnaie	109
9. Les lacunes dans les images	110
10. Le nombre de doigts	113
11. Compter à reculons de 20 à 1	114
12. Le coût de timbres	116

TABLE DES MATIÈRES

Chapitre	page
III.- ETUDE CRITIQUE ET COMPARATIVE DES ITEMS PARTICULIERS (suite)	
13. Ecrire une dictée	118
14. Faire des phrases	119
15. Trouver des rimes	124
16. Disposer des poids	127
17. Les absurdités	129
18. Nommer des mots en une minute	132
19. Le plan de recherche	135
20. Les phrases en désordre	137
21. Le problème d'induction	139
22. Renverser les aiguilles de l'horloge	141
23. Les problèmes d'ingéniosité	145
24. La compréhension de la lecture	146
25. Le dessin au miroir	148
IV.- ETUDE CRITIQUE DU TEST EN GENERAL	149
CONCLUSION	166
BIBLIOGRAPHIE	169
Appendice	
1. <u>THE CANADIAN INTELLIGENCE EXAMINATION - FOREWORD</u>	171
2. <u>THE USE OF THE CANADIAN INTELLIGENCE TEST</u>	177
3. <u>UNE FORMULE D'EXAMEN DU CANADIAN INTELLIGENCE EXAMINATION</u>	184
4. <u>ABSTRACT OF Une étude critique et comparative du Canadian Intelligence Examination</u>	185

LISTE DES TABLEAUX

Tableau	page
<p>I.- <u>Sous-tests du Stanford-Binet</u> conservés par Amoss dans l'organisation de son <u>C.I.E.</u></p>	14
<p>II.- <u>Sous-tests du Stanford-Binet</u> éliminés par Amoss dans l'organisation de son <u>C.I.E.</u></p>	15
<p>III.- <u>Sous-tests ajoutés</u> par Amoss à la révision <u>Stanford-</u> <u>Binet</u> dans l'organisation de son <u>C.I.E.</u></p>	17
<p>IV.- Les séries de chiffres répétés à la suite employés par divers auteurs aux différents âges</p>	56
<p>V.- Les séries de chiffres répétés à reculons employés par divers auteurs aux différents âges</p>	59a
<p>VI.- Les absurdités proposées par différents auteurs à dix ans</p>	131

LISTE DES FIGURES

Figure	page
1. La série de dessins à copier proposée par Amoss dans son <u>C.I.E.</u>	26
2. Les plis à faire dans la première partie du test de pliage de feuilles de Kuhlmann à six ans	36
3. Les plis à faire dans la seconde partie du test de pliage de feuilles de Kuhlmann à six ans	37
4. Les dessins à reproduire de mémoire à dix ans	88
5. Les dessins à reproduire de mémoire, proposés par Amoss, à huit ans	91
6. Dessin à reproduire de mémoire, proposé par Amoss à adulte moyen	91
6a. Position des triangles préconisée par Binet pour le jeu de patience	103
7. Position des triangles préconisée par Terman pour le jeu de patience	104
8. Position des triangles préconisée par Amoss pour le jeu de patience	104

INTRODUCTION

Le test d'intelligence tel que nous le connaissons aujourd'hui est de développement récent. Comme tant d'autres résultats de la science moderne, on peut dire que le test d'intelligence est apparu comme la réalisation d'un besoin qui existait. La science de la psychologie avait fait assez de progrès dans le domaine du test mental pour être prête à remplir ce besoin quand le temps vint.

Ce fut en 1905 qu'Alfred Binet en collaboration avec Thomas Simon, présenta sa fameuse échelle métrique d'intelligence qu'il fit suivre par deux révisions en 1908 et en 1911. Encore aujourd'hui on travaille sur le concept d'intelligence de Binet et sur ses échelles métriques. Les psychologues reconnurent vite la valeur indiscutable de ces échelles et ils en firent des traductions, des adaptations et des remaniements. Notons en particulier les travaux de Bobertag en Allemagne, de Decroly et Degand en Belgique, de Descoedres en Suisse, de Trèves et Saffiotti en Italie, de Burt en Angleterre, et enfin de Goddard, de Terman et de Kuhlmann aux Etats-Unis. C'est en France, son pays d'origine, que l'échelle métrique connut le moins de succès.

La révision qui connut le plus de vogue fut sans contredit le Stanford-Binet¹ de Terman publié en 1916. On lui reconnut une plus grande utilité qu'on ne l'avait tout d'abord prévu. En présence des services nombreux et importants que ce test était appelé à rendre, Terman crut

1 Lewis M. Terman, The Measurement of Intelligence, Boston, Houghton-Mifflin, 1916, pp. 362.

bon de faire une nouvelle révision aussi parfaite que possible, le New Revised Stanford-Binet². Cette dernière révision demanda à Terman dix longues années de travail en collaboration avec Merrill et avec l'aide de plusieurs de ses gradués de l'Université Stanford. Contrairement à la conception fautive assez courante parmi les éducateurs qu'il est facile de construire un test d'intelligence, cette tâche exige une grande habileté technique, une patience à toute épreuve et une somme de travail énorme

En 1940, le Dr Amoss en collaboration avec le Dr C.G. Stogdill présentait le Canadian Intelligence Examination³. Cette publication était la première du genre produite au Canada par un chercheur canadien. Quatre années plus tôt le Dr Amoss avait présenté le Ontario School Ability Examination⁴, un test de performance destiné à être utilisé avec des sujets qui sont sourds ou qui ne connaissent pas la langue anglaise. C'est le premier test qui a retenu notre attention: il fera l'objet de la présente étude.

Plusieurs raisons nous ont incité à faire cette étude. Nous croyons qu'il est fort à propos de faire une étude critique d'un test d'intelligence qui sert dans les écoles d'Ontario. Cette étude s'impose

2 Lewis M. Terman et Maud A. Merrill, Measuring Intelligence, Boston, Houghton-Mifflin, 1937, pp. 461.

3 Harry E. Amoss et C.G. Stogdill, The Canadian Intelligence Examination, Toronto, The Ryerson Press, 1940, pp. 69.

4 Harry E. Amoss, The Ontario School Ability Examination, Toronto, The Ryerson Press, 1936, pp. 54.

d'autant plus que le test lui-même n'a pas encore reçu grand contrôle expérimental. D'autre part, cette recherche est possible. Nous croyons avoir les instruments, la direction et la préparation nécessaires à une pareille entreprise.

Nous avons choisi un test d'intelligence générale à cause de notre vif intérêt pour ce genre de tests. Nous entrevoyions tout d'abord la possibilité d'une adaptation française de ce test pour la population canadienne-française. Une fois nos études commencées, nous n'avons pas jugé que le test en valait la peine. Nous nous bornerons donc à faire une étude critique et comparative du Canadian Intelligence Examination. Dans notre étude nous le comparerons souvent avec les tests d'auteurs aussi connus que Terman et Kuhlmann.

Le rapport de cette recherche débutera par la présentation du Canadian Intelligence Examination selon Amoss. Un deuxième chapitre sera consacré à l'étude de chaque série d'items à cinq points de vue différents: 1° leurs sources, 2° les modifications apportées, 3° les directives, 4° le cotage proposé, et 5° un jugement sur la valeur des items. Le troisième chapitre traitera les items particuliers qui ne se présentent qu'une fois; ils seront étudiés aux mêmes points de vue que les séries d'items. Un quatrième chapitre terminera ce travail par une étude critique et comparative du test en général: l'allure générale, le matériel, la méthodologie et surtout les qualités métrologiques du test et la technique d'administration seront les principaux points étudiés dans cette dernière partie.

CHAPITRE I

LE CANADIAN INTELLIGENCE EXAMINATION SELON AMOSS

Le Dr Harry Amoss¹ présente son Canadian Intelligence Examination dans la préface² de son manuel. Nous y trouvons 1^o les raisons qui l'ont poussé à publier son test, 2^o le but qu'il se proposait, 3^o sa conception de l'intelligence, 4^o son adhésion à la révision Stanford de l'examen d'intelligence de Binet-Simon, et 5^o les principales modifications qu'il a apportées à cette même révision. Nous allons relever ces cinq points pour les élaborer au besoin, et y ajouter même une petite note critique.

1. La nécessité de ce test individuel.

La révision Stanford-Binet³ servait déjà depuis plus de vingt ans à la branche des classes auxiliaires du Ministère de l'Éducation d'Ontario pour les élèves qui méritaient une attention spéciale. En 1937, la publication du Revised Stanford-Binet⁴ par Terman fit croire à Amoss que la révision du Stanford-Binet serait immédiatement discontinuée et qu'ainsi les instituteurs seraient privés d'un examen

1 Harry Amoss et Charles Stogdill, Canadian Intelligence Examination, Toronto, Ryerson Press, 1940, p. 5. Le Dr Amoss B.A., D. Paed., est Director of Professional Training au Département de l'Éducation de l'Ontario; il a partagé son droit d'auteur avec Charles Stogdill.

2 On peut lire une copie de cette préface à l'Appendice 1.

3 Lewis M. Terman, The Measurement of Intelligence, Boston, Houghton Mifflin Co., 1916.

4 Lewis M. Terman et Maud A. Merrill, Measuring Intelligence, Boston, Houghton Mifflin Co., 1937.

simple, efficace et peu coûteux. Il crut donc le moment venu d'offrir aux instituteurs un examen assez sérieux qui lui paraissait mieux adapté, à la fois aux milieux ruraux, et aux milieux urbains.

Le docteur n'avait pourtant rien à craindre à ce sujet. Le C.I.E. fut publié en 1940, et jusqu'à date le Stanford-Binet de 1916 est encore sur le marché. Les éditeurs Houghton-Mifflin Co. de Boston nous avertissent⁵ que le test se vend toujours et qu'ils ont suffisamment de matériel et de manuels pour répondre aux demandes ordinaires pendant au moins un an. Il leur est toutefois impossible de garantir une nouvelle impression.

L'auteur attache quand même beaucoup d'importance à son projet qui est d'offrir aux instituteurs un examen facile à donner, simple et peu coûteux. Il semble même qu'à ses yeux le Revised Stanford-Binet de 1937 ne serait pas doté de ces précieuses qualités. Voyons plutôt.

Il est vrai que le Revised Stanford-Binet coûte beaucoup plus cher que le C.I.E. On peut se procurer le matériel et le manuel de ce dernier chez Ryerson Press à Toronto pour la somme minime de \$2.00, tandis que le matériel et le manuel du Revised Stanford-Binet coûtent environ \$15.00 aux Canadiens. Sans doute, c'est une forte dépense, mais ce n'en reste pas moins au point bien secondaire pour le technicien en quête des meilleurs instruments qui soient pour la mesure

⁵ Correspondance personnelle de l'auteur, lettre de Houghton Mifflin Co. datée du 8 déc. 1943.

de l'acuité et de la maturité de l'intelligence. L'instituteur hésitera peut-être longtemps devant cette dépense, mais quand il aura déboursé plusieurs fois le même montant pour des cours qui le prépareront à se servir du test, il sera fort aise de posséder un meilleur instrument.

En deuxième lieu, l'auteur trouve le Revised Stanford-Binet trop élaboré. Son propre test l'est-il moins? même s'il n'a pas deux échelles? D'ailleurs point n'est besoin de maîtriser les deux échelles si on les trouve trop longues! Vingt années d'expérience prouvèrent à Terman l'insuffisance du Stanford-Binet. Le test souffrait de deux grands défauts: les sous-tests des bas-âges étaient trop faciles tandis que les sous-tests des âges plus avancés étaient trop difficiles et trop peu nombreux. Le New-Revised Stanford-Binet remédia à ces deux lacunes. Tant d'années d'expérience produisirent un test plus objectif, mieux gradué et facilement utilisable par des personnes de peu d'expérience.

En dernier lieu, l'auteur se propose d'offrir un test facile à donner. A-t-il vraiment réussi? Quand nous avons donné le C.I.E. et le Revised Stanford-Binet, ce dernier nous a paru sans contredit plus facile à donner et aussi plus facile à interpréter parce que plus objectif. Il laisse les choses moins au hasard, les directives sont plus claires, le cotage plus facile et la classification des réponses acceptables mieux définie. Par contre dans le C.I.E. il peut se glisser beaucoup plus de subjectivité. Ce point particulier sera traité au long dans un chapitre subséquent.

Une autre raison qui induisit l'auteur à publier son test fut la constatation depuis quelques années d'une tendance de plus en plus marquée parmi les gens, même les plus avertis, à confondre les fonctions d'un test de groupe et d'un test individuel. On employait trop facilement le premier quand on aurait dû administrer l'examen individuel. C'était là, pensait-il, une autre bonne raison pour présenter son test.

En effet, les autorités en la matière reconnaissent bien comme l'auteur que l'usage d'un test de groupe doit se limiter à l'étude du groupe en question. C'est un moyen précieux d'obtenir pour fins de comparaisons une idée générale du groupe, de la classe, de l'école ou de la population étudiée. Les résultats que l'on obtient ainsi valent rarement pour les individus dans le groupe; ils peuvent même nous induire en erreur dans le cas de sujets nerveux ou lents à s'exprimer. C'est dire que l'on ne devrait se servir que d'examens individuels quand il s'agit d'étudier, de classer ou de promouvoir des élèves. Le moment semblait donc opportun de mettre entre les mains des techniciens et des instituteurs un test d'intelligence individuel qui serait assez simple pour se donner avec succès après une courte période d'apprentissage.

2. Le but que se proposait l'auteur.

Le C.I.E. a été préparé pour mesurer l'intelligence native ou acquise. L'auteur a-t-il pleinement réalisé son but? Cet instrument mesure-t-il vraiment ce qu'il prétend mesurer? C'est tout le problème de la justesse ou de la validité de son test.

Amoss propose son instrument aux instituteurs de la province d'Ontario, et pour la mesure de l'intelligence, et surtout de la scolapitude des enfants ordinaires, des enfants brillants et des cas frontières de plus de quatre ans.

Dans un chapitre subséquent, nous aurons l'occasion de mettre à l'épreuve les prétentions de l'auteur.

3. La conception de l'intelligence selon Amoss.

L'auteur explique que celui qui possède à un haut degré la capacité d'organiser les moyens en vue d'une fin déterminée en évitant le tâtonnement, celui-là est doué d'une intelligence supérieure, qu'il s'agisse d'écrire une composition littéraire, de creuser un fossé, de conduire des troupes à la victoire ou d'organiser une fête quelconque. Une personne d'intelligence inférieure s'en tirera tant bien que mal ou ne s'en tirera pas du tout parce qu'elle est incapable, 1^o de tenir compte de tous les faits se rapportant à une situation, 2^o de tenir présent à l'esprit des données précises tandis qu'elle formule son plan d'action, et 3^o de rassembler et d'organiser les détails pour la bonne exécution d'un plan. Ainsi un manufacturier négligera de tenir compte de la possibilité que des troubles politiques en Asie peuvent lui couper ses provisions de tungstène; un élève oubliera d'inclure les trois jours de grâce quand il résoud un problème d'escompte; et un mineur qui creuse un puits omettra de calculer le sondage d'une veine faillée.

L'auteur en vient donc à conclure qu'on peut définir l'intelligence d'une façon négative comme la capacité d'éviter le tâtonnement, et d'une façon positive comme la capacité de comprendre les faits qui se rapportent à un ensemble de circonstances et de s'en servir comme moyens pour arriver à une fin.

Intelligence, therefore, may be defined negatively as ability to eliminate trial and error forms of behaviour, and positively as ability to grasp the facts of a situation and use these to formulate plans adequate to occasion. In brief, intelligence is ability to plan ⁶.

Tout en nous faisant connaître sa définition de l'intelligence, Amoss nous fait part de sa façon de concevoir les mesures de l'intelligence. Quelques lignes plus loin on lit:

Intelligence testing will only become an exact science when there shall have been evolved a measuring unit formulated in terms of ability to establish certain mathematically defined relations among certain mathematically defined data, and which yard stick shall be as applicable in estimating the size of a job as in determining the capability of the worker ⁷.

C'est tout le problème du nombre en psychologie qu'il soulève d'un simple coup de main. Amoss voit bien que le nombre obtenu dans un test n'est encore qu'un symbole; dans son ardeur de savant il le voudrait voir devenir une mesure véritable où l'unité de grandeur donnerait prise aux relations mathématiques exactes. Cette unité mathématique mesurerait tant la tâche à accomplir que le travailleur qui doit l'accomplir.

⁶ Amoss, op. cit., p. 7.

⁷ loc. cit.

Hélas, nous croyons que les données psychologiques, essentiellement qualitatives ne se laisseront jamais caser dans les unités d'aptitudes réservées par le bon Dr Amoss.

Malgré notre peu d'espoir de voir se réaliser les fervents désirs du savant, nous admettons avec lui, que tels quels les tests d'intelligence peuvent rendre de grands services. Comme instrument de diagnostic, et même de pronostic, ils ont fait leur preuve à la fois en éducation et en hygiène mentale. Les dix ans d'expérience de l'auteur pendant lesquels il a pu examiner des milliers d'enfants et contrôler même les suites de ses conseils, confirment bien cette conclusion.

4. Fidélité à la révision Stanford-Binet.

Le Dr Amoss nous dit dans la préface de son manuel qu'il s'en est tenu à la structure générale du Stanford-Binet. C'est vrai: tests, sous-tests, âges mentaux, etc. suivent de près ceux de Terman. Comme la première échelle de Terman, celle d'Amoss s'étend de trois ans à adulte supérieur. De trois à dix ans, chaque année comprend six sous-tests auxquels on accorde deux mois de crédit; comme dans cette première échelle il n'y a pas de tests à onze ans et à treize ans, mais à douze ans il y a huit tests auxquels sont attribués trois mois de crédit et à quatorze ans, six tests de quatre mois de crédit. On trouve une légère variante aux tests d'adulte moyen qu'Amoss appelle les tests de XIV-XVII ans. Terman accorde cinq mois de crédit à chacun des six tests tandis qu'Amoss en accorde six. Aux tests d'adulte supérieur qu'Amoss intitule tests de

XVII-XX ans, on trouve six tests avec un crédit de six mois chacun comme le fait d'ailleurs Terman.

L'âge mental se calcule de la même façon que dans le Stanford-Binet: à l'âge de base sont ajoutés tous les crédits des tests réussis et ce total en années et en mois donne l'âge mental du sujet.

Le calcul des quotients intellectuels se fait de la même façon dans les deux tests: c'est-à-dire en divisant l'âge mental que le sujet a obtenu sur le test par son âge chronologique et en multipliant par cent. Pour trouver le quotient intellectuel d'adultes, Terman propose un diviseur de 192 mois (seize ans) dans le Stanford-Binet parce que, selon lui, le développement intellectuel cesse à cet âge. Il changera cette façon de trouver le quotient intellectuel dans sa nouvelle révision. Au lieu de faire cesser le développement intellectuel tout d'un coup à seize ans, il propose l'hypothèse que le ralentissement commence à treize ans pour cesser complètement à seize ans. Pour trouver le quotient intellectuel des sujets de treize à seize ans, l'âge chronologique du sujet sera calculé comme treize ans plus deux tiers des mois additionnels qu'il aura vécu puisqu'il considère que ce ralentissement se fait à raison de quatre mois par année. Ainsi l'âge chronologique d'un sujet de quatorze ans sera considéré comme treize ans et huit mois, un âge chronologique de quinze ans comme quatorze ans et quatre mois, et un âge chronologique de seize ans comme quinze ans qui sera le diviseur le plus élevé pour le calcul du quotient intellectuel.

Amoss fait un compromis entre ces deux méthodes. Au lieu d'imiter le Stanford-Binet comme il le fait dans tout le reste, il choisit quinze ans d'âge chronologique (180 mois) comme diviseur maximum pour le calcul des quotients intellectuels des sujets au-dessus de cet âge. Les quotients intellectuels de sujets de plus de quinze ans d'âge chronologique obtenus sur le C.I.E. seront donc plus élevés que ceux du Stanford-Binet puisque le dénominateur est de 180 contrairement à 192 qui est le dénominateur du Stanford-Binet; les quotients intellectuels obtenus sur le C.I.E. à quatorze et à quinze ans seront plus bas que ceux du New Revised Stanford-Binet puisqu'Amoss ne tient pas compte du ralentissement du développement intellectuel à ses âges. Les quotients intellectuels de sujets de treize ans ou plus obtenus sur le C.I.E. ne seront pas directement comparables aux quotients obtenus sur une ou l'autre des échelles de Terman. Bien que des quotients intellectuels soient rarement comparables, voire même jamais, au premier abord on serait porté à croire qu'ils le sont puisqu'Amoss a imité de très près le Stanford-Binet.

5. Principales modifications apportées à la révision Stanford-Binet.

Les principaux changements que l'auteur prétend avoir fait subir aux tests de la révision Stanford-Binet sont les suivants:

1° Il a simplifié le langage des directives dans la présentation des sous-tests. Voici ce que Terman dit dans Measuring Intelligence à

propos des directives. Nous citons:

It cannot be too strongly emphasized that unless we follow a standardized procedure, the tests lose their significance. Sometimes the omission or the addition of a single phrase in giving the test will alter materially the significance of the response. Only the trained psychologist can vary the formula without risk of invalidating the result, and even he, must be on guard ⁸.

Si l'auteur a simplifié les directives, il se peut fort bien qu'il ait aussi changé la valeur des items, et un item, v.g. placé au niveau de six ans devienne dès lors un test plus facile, ou peut-être même plus difficile. Quelle expérimentation l'auteur a-t-il fait pour savoir où placer l'item ainsi changé dans son échelle? A l'encontre de Terman qui consacre quatre chapitres dans Measuring Intelligence et sept chapitres dans The Measurement of Intelligence à expliquer sa méthode de standardisation, Amoss n'en souffle guère mot. Craignait-il ennuyer ses lecteurs avec des explications pourtant essentielles à un travail scientifique?

2° Il dit qu'il s'est efforcé de définir plus rigoureusement les réponses acceptables. Pour un petit nombre des tests, oui. Mais même pour ceux-là il aurait fort bien pu aider l'expérimentateur à formuler son jugement par quelques dessins modèles. Dans les tests où le sujet doit faire un dessin ou en reproduire un de mémoire, Terman consacre une page entière à des illustrations qui servent d'exemple pour la classification des réponses bonnes ou mauvaises.

8 Terman, Measuring Intelligence, op. cit., p. 131.

Amoss aurait dû dresser des listes de réponses acceptables ou insuffisantes. Ainsi on trouve six tests de ce genre dans le C.I.E. Des exemples aideraient beaucoup l'expérimentateur, surtout dans les tests des âges avancés où on demande la différence entre des mots qui, à première vue semblent être synonymes, v.g. la différence entre l'espérance et l'attente, le calme et la patience, le probable et le possible, la noblesse et la dignité, le savoir et la sagesse, avoir et posséder.

Ce que nous venons de dire plus haut s'applique aussi au test de vocabulaire. L'auteur présente une liste de cinquante-cinq mots et c'est tout. Encore ici faut-il se débrouiller comme bon nous semble. Par contre, Terman ne donne jamais moins de cinq illustrations de bonnes ou de mauvaises réponses pour chaque mot des vocabulaires présentés dans le Revised Stanford-Binet.

Dans le genre de tests ci-haut mentionnés, la variété des réponses peut être aussi grande que le nombre de sujets, et il est absolument certain que les examinateurs coteront ces réponses d'une manière différente. Il est évident que toutes les réponses bonnes ou mauvaises ne peuvent être notées mais un certain nombre d'illustrations servant de modèles aideraient beaucoup pour que le cotage se fasse avec objectivité.

Obtenir plus d'objectivité dans l'appréciation des réponses fut un des principaux buts de Terman dans le New Revised Stanford-Binet. Quand il a fait passer les tests pour la sélection finale des items, il a exigé que les examinateurs notent toutes les réponses textuellement. Avant de déterminer ses deux échelles finales de la standardisation des

données obtenues il a trouvé nécessaire de vérifier et de coter de nouveau toutes les formules des tests. Bien que sept examinateurs aient coté les réponses provisoirement à mesure que les tests étaient donnés, il était inévitable qu'il se présente des variantes dans l'application des règles données. En faisant passer le test à plus de six mille sujets, des situations inattendues surgissaient et des réponses imprévues se rencontraient. Le cotage fut donc refait par deux personnes sous la direction de l'auteur et à mesure qu'elles cotaient, elles dressaient des listes de bonnes et de mauvaises réponses pour chaque item. Comme résultat on a pu proposer des principes et des illustrations bien définis partout où l'examineur doit se servir de son jugement dans l'évaluation des réponses. La facilité et l'objectivité du cotage ont joué un rôle décisif dans l'acceptation ou le refus d'un item dans la sélection finale des deux échelles.

Si Amoss avait dressé de pareilles listes pour chaque item qui en a besoin, son test y aurait beaucoup gagné en objectivité; mais, tel quel, il n'y a pas deux personnes qui obtiendraient le même résultat en cotant le même test, car trop de subjectivité peut s'y glisser. L'auteur qui se proposait de définir plus rigoureusement les réponses acceptables n'a guère réalisé ses projets.

3^o Amoss nous parle ensuite des sous-tests du Stanford-Binet qu'il a omis dans son échelle. Avant de considérer ceux-ci, voyons un peu ceux qu'il a conservés. Comme il le disait dans la section précédente il a suivi de près la révision Stanford-Binet. Nous irons plus loin et nous

dirons qu'il l'a imitée et même qu'il l'a presque copiée. En effet, si nous consultons le Tableau I nous voyons une longue liste de sous-tests, soixante et un en fait, qu'il a pris presque textuellement du Stanford-Binet pour les présenter dans le C.I.E. Il est vrai qu'il leur a fait subir quelques petites modifications que nous verrons dans les deux chapitres subséquents, mais est-ce ainsi que l'on fait quand on a la prétention d'offrir un test nouveau?

Des soixante tests qu'Amoss a gardés dans le C.I.E., Terman en a éliminé dix-sept dans le New Revised Stanford-Binet à cause de leur peu de validité, de la difficulté à les coter et du danger d'aider l'élève. Dans le même Tableau I ils sont précédés d'un apostrophe.

Amoss n'a pas osé parler des sous-tests qu'il avait conservés car c'est le grand nombre, mais il prend soin de souligner qu'il en a omis un bon nombre. Nous les trouvons au Tableau II, vingt-six en fait. La raison? il croit que le hasard y joue une trop grande part. De ces vingt-six sous-tests éliminés par Amoss, Terman n'en supprime que quatorze dans sa révision de 1937; dans le même Tableau II ils sont précédés d'un apostrophe.

D'après ces deux tableaux, quatorze sous-tests ont été éliminés par nos deux auteurs, douze ont été gardés par Terman et rejetés par Amoss et dix-sept ont été conservés par Amoss et omis par Terman. Il y a donc trente et un sous-tests sur lesquels Terman et Amoss diffèrent d'opinion. A qui donnerons-nous raison? A Amoss? mais sur quoi fonder notre jugement? Il affirme toujours sans ne jamais rien prouver. A Terman alors? Nous serions bien portés à le faire, puisqu'il ne néglige pas de justifier son choix dans les quatre premiers chapitres de Measuring Intelligence.

Tableau I.-

Sous-test¹ du Stanford-Binet conservés par Amoss dans l'organisation de son C.I.E.

Nom des sous-tests	Année et n°	Page ¹
Nommer des objets	III, 2	56
Énumération d'objets dans les images	III, 5	145
Répétition de trois chiffres	III, supp.	150
*Répétition de syllabes	III, 6	149
Copier un carré	IV, 4	155
Compter trois sous	IV, 3	154
Discrimination des formes	IV, 2	152
Compréhension	IV, 5	157
Répétition de quatre chiffres	IV, 6	159
Répétitions de syllabes	IV, supp.	160
*Nommer des couleurs	V, 2	163
Définition par l'usage	V, 4	167
Jeu de patience	V, 5	169
Trois commissions	V, 6	172
*Distinguer la gauche de la droite	VI, 1	175
*Reconnaître des pièces de monnaie	VI, 5	184
Compter treize sous	VI, 3	180
Lacune dans des images	VI, 2	178
*Répétition de syllabes	VI, 6	185
Copier un losange	VII, 6	204
Nombre de doigts	VII, 1	189
*Description d'images	VII, 2	190
Répétition de cinq chiffres	VII, 3	193
*Différences	VII, 5	199
Répétition de trois chiffres à reculons	VII, supp.	207
Les jours de la semaine	VII, supp.	205
*Compter à reculons de 20 à 1	VIII, 2	213
Ressemblances	VIII, 4	217
*Coût de timbres	IX, supp.	253
Vocabulaire	VIII, 6	224
*Écrire une dictée	VIII, supp.	231
Répétition de quatre chiffres à reculons. . . .	IX, 4	242
*Construction de phrases	IX, 5	242
Trouver des rimes	IX, 6	248
*Les mois de l'année	IX, supp.	251
*Les poids	IX, 2	236
Vocabulaire	X, 1	255
Reproduire un dessin de mémoire	X, 3	260
Absurdités	X, 2	255
Nommer des mots en une minute	X, 6	272
Répétition de six chiffres	X, supp.	277
Vocabulaire	XII, 1	281
Répétition de cinq chiffres à reculons	XII, 6	301
La balle et le champ	XII, 3	286
Définitions	XII, 2	281
Phrases mélangées	XII, 4	286
Interprétation d'images	XII, 7	302
Ressemblance entre trois choses	XII, 8	306
Vocabulaire	XIV, 1	310
Problèmes d'induction	XIV, 2	310
*Répétition de sept chiffres	XIV, supp.	322
Raisonnement mathématique	XIV, 5	319
*Renverser les aiguilles de l'horloge	XIV, 6	321
*Différences	XIV, 3	313
Vocabulaire	Adulte moyen, 1	324
Différences	Adulte moyen, 3	324
Répétition de six chiffres à reculons	Adulte moyen, 5	327
Raisonnement mathématique	Adulte moyen, 4	327
Problèmes d'ingéniosité	Adulte sup., 6	345
*Répétition de sept chiffres à reculons	Adulte sup., 5	345

{*} Set apostrophe devant les sous-tests indique ceux que Terman a éliminés dans le New Revised Stanford-Binet; il y en a dix-sept en tout.

TABLEAU II.-

Sous-tests du Stanford-Binet 1916 éliminés par Amoss dans l'organisation de son C.I.E.

Nom des sous-tests	Année et n°	Page ¹
Indiquer les parties du corps	III, 1	142
*Donner le sexe	III, 4	146
*Donner le nom de famille	III, 5	147
*Comparaison de lignes	IV, 1	151
Comparaison esthétique	V, 3	165
*Donner l'âge	V, 7	173
Compréhension II	VI, 4	181
*Avant-midi et après-midi	VI, 7	187
*Faire une boucle	VII, 4	196
Compréhension III	VIII, 3	215
*Les définitions	VIII, 5	221
*Nommer six pièces de monnaie	VIII, 7	231
*Donner la date du mois	IX, 1	234
Donner huit idées dans une lecture	X, 4	262
*Compréhension IV	X, 5	268
Répéter des phrases de mémoire	X, 8	277
*La planchette à encastrement	X, 9	278
*L'interprétation de fables	XII, 5	290
L'interprétation de faits	XIV, 4	315
*L'interprétation de fables	Adulte moyen, 2	324
Les codes	Adulte moyen, 6	330
Répéter des phrases de mémoire	Adulte moyen, 7	332
*Compréhension de relations physiques	Adulte moyen, 8	333
Le découpage de Binet	Adulte sup., 2	338
Répéter huit chiffres	Adulte sup., 3	340
Répéter les idées contenues dans une lecture	Adulte sup., 4	340

(*) Cet apostrophe devant le sous-test indique ceux-là que Terman lui-même a éliminés; il y en a quatorze en tout.

1 Lewis M. Terman, The Measurement of Intelligence, Boston, Houghton-Mifflin Co., 1916.

Par ailleurs, Amoss a fait des additions, ce sont les tests que nous trouvons au Tableau III. Pour suppléer aux tests du Stanford-Binet qu'il a omis, il en ajoute dix-huit qui, à l'exception des tests de pliage de feuilles qui ont été imaginés par Kuhlmann, peuvent lui être attribués. Sur un total de soixante-dix-huit tests c'est une contribution assez mince.

4. Le Dr Amoss nous explique ensuite comment il a choisi les cinquante-cinq mots de son test de vocabulaire. Son but principal est d'éviter des expressions locales ou littéraires. Il choisit donc ses mots du langage courant des centres urbains et ruraux plutôt que dans un dictionnaire.

Avant d'introduire un vocabulaire dans un examen comme celui-ci, il faudrait le standardiser avec soin. Autrement que vaut-il? Or, pareille standardisation exige une recherche considérable. Amoss s'est-il donné cette peine? Ou bien a-t-il laissé le hasard guider son choix? Autant de questions toujours laissées sans réponse par son manuel.

5. Pour donner moins d'importance à l'aptitude verbale l'auteur a introduit plus de calcul mathématique. Ces problèmes d'arithmétique et ces séries mathématiques que nous retrouvons à quatre endroits différents remédient-ils vraiment à ce défaut qu'il croit exister? Encore ici Amoss affirme sans grande preuve. Terman, au contraire, justifie l'introduction de tout item.

6. L'auteur nous avertit ensuite qu'il a accordé un certain crédit à des réponses partielles pour obtenir un résultat plus finement

Tableau III.-

Sous-tests ajoutés par Amoss à la révision Stanford-Binet 1916, dans l'organisation de son C.I.E.

Nom des sous-tests	Année et n°	Page ¹
Copier le cercle	II-III, 1	32
Copier la croix	II-III, 2	33
Le pliage de feuilles	II-III, 3	33
Le pliage de feuilles	III-IV, 3	36
Copier le triangle	IV-V, 1	38
Le pliage de feuilles	IV-V, 3	38
Copier le pentagone	V-VI, 2	40
Copier l'étoile	VII-VIII, 1	46
Dessin reproduit de mémoire	VII-VIII, 6	49
Faire de la monnaie	VIII-IX, 1	50
Les définitions	VIII-IX, 2	51
Problèmes d'arithmétique	IX-X, 2	54
Problèmes d'arithmétique	X-XII, 2	58
Dessin de mémoire	XIV-XVII, 2	65
Séries mathématiques	XIV-XVII, 6	67
Les différences	XVII-XX, 3	68
Compréhension de la lecture	XVII-XX, 5	68
Dessin au miroir	XVII-XX, 6	69

¹ Harry Amoss et Charles Stogdill, Canadian Intelligence Examination, Toronto, Ryerson Press, 1940.

gradu  et pour  liminer une trop grande indulgence dans l'interpr tation des r ponses. Terman se servait de la m thode "all or none" pour coter. Selon ce proc d  le sujet doit bien r pondre   un certain nombre d'items pour obtenir un cr dit dans le sous-test en question., v.g. dans le cas des d finitions   neuf ans, Terman exige trois bonnes d finitions sur quatre pour accorder les points qui sont deux mois d' ge mental. Amoss cote cet item de la m me fa on mais il donne en plus un mois d' ge mental pour deux bonnes d finitions. Il pr tend arriver ainsi   des r sultats plus finement gradu s. Si l'item est bien plac  dans l' chelle de l'un ou de l'autre, les r sultats seront les m mes. Ce qui importe c'est qu'Amoss n'aille pas placer dans son  chelle un sous-test auquel il accorde ce cr dit partiel au m me endroit qu'il occupait dans celle de Terman, car il n'a plus la m me valeur et ne serait pas valide pour l' ge en question.

Un exemple illustrera le point. Les directives de Terman et celles d'Amoss sont identiques et les figures sont les m mes dans le sous-test des formes g om triques   quatre ans. Terman accorde deux mois d' ge mental pour sept bons choix; pour moins de sept il n'accorde rien. Amoss donne deux mois d' ge mental pour sept bons choix et un mois pour cinq bons choix. Donc, au m me niveau, avec un test identique un sujet qui ne donne que cinq bonnes r ponses se verra accorder un mois d' ge mental de plus sur le C.I.E. que sur le Revised Stanford-Binet. Si plusieurs items du C.I.E. sont cot s ainsi le r sultat final donnera un  ge mental plus  lev ; par suite les deux r sultats ne sont plus comparables comme Amoss semble vouloir le laisser croire.

7. Comme dernière modification l'auteur nous dit qu'il a fixé l'âge moyen de la maturité mentale à quinze ans plutôt qu'à seize. Il a opéré ce changement parce qu'il a constaté que presque tous les sujets soumis au test entre huit et treize ans obtenaient un quotient intellectuel beaucoup plus bas quand ils subissaient l'épreuve une deuxième fois à seize ans ou dix-sept ans. L'usage de 180 comme dénominateur maximum semblait rectifier en grande partie cette dénomination et tendre à garder le quotient intellectuel constant.

Terman a résolu ce problème d'une façon différente dans le New Revised Stanford-Binet. Il a trouvé que le quotient intellectuel diminuait quand on donnait le test au même sujet à l'âge de seize ou dix-sept ans parce que les tests des âges plus avancés étaient trop peu nombreux et trop difficiles. Alors il a introduit un plus grand nombre de tests standardisés avec soin pour les âges plus avancés. Terman nous prévient que les quotients intellectuels sont problématiques après l'âge de quinze ans puisqu'on atteindrait la maturité mentale à cet âge. C'est pourquoi on ne doit se servir de pareils quotients qu'avec beaucoup de discrétion et alors seulement comme indices de catégories comme celles de l'idiotie, l'imbécillité, les cas frontières, l'intelligence normale, la supérieure ou très supérieure.

En résumé, nous avons vu dans ce premier chapitre: 1° les raisons qu'invoque Amoss afin de prouver la nécessité de présenter un nouveau test d'intelligence, raisons d'ailleurs assez peu fondées; 2° le but qu'il se proposait d'atteindre: offrir un test moins coûteux et moins volumineux

que le New Revised Stanford-Binet; 3° sa conception de l'intelligence; et 4° son imitation de la structure général du Stanford-Binet; et 5° les principales modifications qu'il a fait subir à ce test, modifications qui se chiffrent au nombre de sept.

CHAPITRE II

ÉTUDE CRITIQUE ET COMPARATIVE DE CHAQUE SÉRIE D'ITEMS

Après ces quelques pages sur le C.I.E. en général, nous nous proposons de faire maintenant une étude critique et comparative de chaque série d'items du test. Dans un chapitre subséquent nous ferons de même pour les items particuliers qui ne se répètent pas aux différents âges. Notre étude portera sur cinq points de vue différents: 1° leurs sources, 2° les modifications apportées, 3° les directives, et 4° le cotage proposé et 5° la valeur des items. Toutes les séries d'items ne comportent pas également une étude détaillée de ces cinq points, et même si la plupart pourraient y être soumis, ils ne le seront pas toujours dans le même ordre.

Avant de passer en revue 1° la série de dessins à copier, 2° la série de pliage de feuilles, 3° les images, 4° les répétitions de chiffres à la suite, 5° la répétition de chiffres à reculons, 6° la répétition de syllabes, 7° les définitions, 8° le test de vocabulaire, 9° les ressemblances et les différences, 10° les problèmes d'arithmétique, 11° les séries mathématiques, 12° compter des sous, 13° les jours de la semaine et les mois de l'année, et 14° les dessins de mémoire, arrêtons-nous quelque peu à examiner le matériel du C.I.E.

Le matériel nécessaire au test est réduit à sa plus simple expression: une montre, un miroir, un canif, un crayon, un jeu de patience, des

sous, quatre images, les cartes qui représentent les dessins à reproduire, les couleurs, les timbres, les formes géométriques et enfin celles sur lesquelles sont écrits les divers problèmes. En somme, tout ce matériel très simple n'offre pas beaucoup d'attrait pour l'enfant. Par contre le matériel du Revised Stanford-Binet bien que très élaboré et plus dispendieux est beaucoup plus intéressant. Il y a une grande poupée en carton, de petites boîtes de carton, un petit chat, un petit chien, une boule, un lit, une locomotive, un batteur d'oeufs, une paire de ciseaux, un dé, une cuillère, des perles, des blocs et un grand nombre d'images sur les différences et les ressemblances, images mutilées et images absurdes; en somme une foule de choses qui, employées dans un grand nombre de sous-tests, intéressent vivement l'enfant et lui font donner son maximum de rendement.

Il faut du concret aux tout petits . Essayer d'éveiller leur intérêt avec de l'abstrait est peine perdue. Terman a bien su mettre à profit sa connaissance de la psychologie de l'enfance sans se préoccuper du prix que pourrait coûter le matériel. Si l'instrument en valait la peine on n'hésiterait pas à délier les cordons de sa bourse! Amoss, lui, lésine! Le résultat? -Médiocre comme on pouvait le prévoir. C'est à se demander si Amoss n'appréciait pas toute la valeur d'un bon instrument.

Des expériences personnelles répétées nous permettent de pareilles affirmations. Nous avons donné le C.I.E. et le New Revised Stanford-Binet aux enfants de bas-âges et il nous a toujours été plus difficile d'intéresser l'enfant avec les sous-tests du C.I.E. Les jeunes enfants

aiment à faire les choses eux-mêmes. Le Revised Stanford-Binet leur en offre l'occasion. Enumérons plutôt: remettre les blocs en place dans la planchette à encastrement, construire une tour ou un pont avec des blocs, enfiler des perles, reproduire une chaîne de perles de mémoire, plier des feuilles, faire une boucle, assortir des boutons, faire le jeu de patience, faire trois commissions, compter des cents, terminer un dessin, reproduire des dessins de mémoire, obéir à de simples commandements, déchiffrer des codes en plus d'identifier une foule d'objets intéressants par des images. Les seuls tests du C.I.E. qui offrent de l'action sont les suivants: copier des dessins, plier des feuilles, compter des cents, faire le jeu de patience, faire trois commissions et reproduire des dessins de mémoire.

Chaque fois que nous avons donné le C.I.E. aux jeunes enfants nous avons toujours commencé par leur donner le test de la construction de blocs de Gesell que l'on trouve dans le Ontario School Ability Examination¹ d'Amoss. Les enfants y prenaient toujours grand plaisir et par conséquent ce sous-test a prouvé d'une grande valeur pour éveiller leur intérêt et gagner leur coopération. Pourquoi l'auteur ne l'a-t-il pas inclus dans son C.I.E. ainsi que d'autres tests du même genre que nous retrouvons dans son Ontario School Ability Examination? Le C.I.E. y aurait beaucoup gagné en intérêt.

1 Harry Amoss, Ontario School Ability Examination, Toronto, Ryerson Press, 1936, pp. 24-27.

Une autre différence entre le C.I.E. et le Revised Stanford-Binet est dans la manière de procéder pour les séries d'items tels que de copier les dessins, de plier les feuilles, de répéter les chiffres à la suite ou à reculons et de répéter les syllabes. Terman les intercale aux différents âges tandis qu'Amoss exige qu'ils soient donnés à la suite jusqu'au point où le sujet manquera un item. Ainsi quand on arrive à la répétition de trois chiffres à l'âge de trois ans, si le sujet réussit cet item, on procédera immédiatement à la répétition de quatre chiffres à l'âge de quatre ans, de cinq chiffres à l'âge de sept ans et ainsi de suite jusqu'au point où l'élève échouera. Cette manière de faire sauve du temps car on n'a pas à répéter les directives au long chaque fois, mais il y a aussi des désavantages que Terman souligne dans Measuring Intelligence. Nous citons:

It is not necessary and often it is not expedient, to give the digit tests of all the different years in succession; that is, without other tests intervening. While this may be permissible with older children, in young children the power of sustained attention is so weak that no single kind of test should occupy more than two or three minutes. Children below six or seven years should ordinarily be given the tests in the order in which they are listed in the record booklet ².

Ce que Terman dit à propos de la répétition de chiffres peut s'appliquer à la répétition de syllabes. Il est évident que faire répéter des séries de chiffres ou de syllabes augmentant en difficulté trois ou quatre fois de suite désintéressera l'enfant et le fatiguera à cause de l'effort soutenu que ce genre de test exige. Le sujet ne donnera certainement

² Lewis M. Terman, The Measurement of Intelligence, Boston, Houghton-Mifflin, 1916, p. 196.

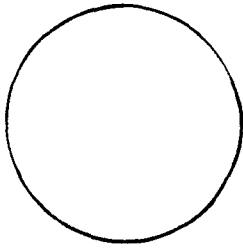
pas un aussi bon rendement que s'ils sont présentés aux différents âges. L'effet contraire peut aussi se produire dans le cas d'un test intéressant comme le pliage de feuille. Ici l'enfant bénéficiera de l'effet de l'apprentissage. En somme il est préférable de ne pas présenter à la fois plusieurs items trop monotones ou trop intéressants car on enlève ainsi au test toute sa variété.

Après ces quelques mots sur le matériel du C.I.E. nous ferons maintenant une étude critique et comparative de chaque série d'items.

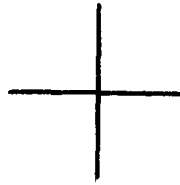
1. La série de dessins à copier.

Dessiner d'après modèle a toujours été un test très satisfaisant aux différents niveaux d'âge où il a été employé. Il est devenu typique dans les échelles d'intelligence. Amoss introduit sept dessins à copier dans le C.I.E.: le cercle et la croix à trois ans, le carré à quatre ans, le triangle à cinq ans, le pentagone à six ans, le losange à sept ans et l'étoile à huit ans.

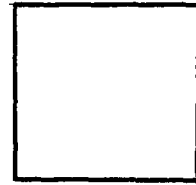
Nous discuterons tout d'abord le cercle, la croix, le carré et le losange, car ce sont les items que les auteurs en la matière ont utilisé le plus souvent.



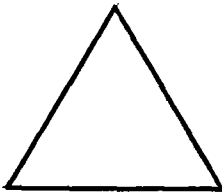
Le cercle



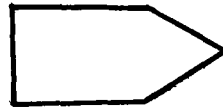
La croix



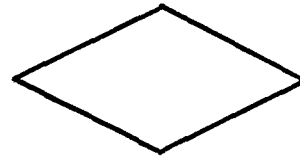
Le carré



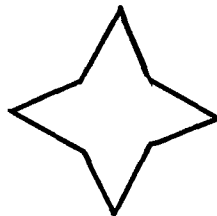
Le triangle



Le pentagone



Le losange



L'étoile

Figure 1. La série de dessins à copier proposés par Amoss dans le C.I.E.

Les sources.-- Le carré et le losange ont d'abord servi à Binet et à Simon³ dans leur échelle métrique, puis à Terman⁴ dans le Stanford-Binet de 1916, et à Kuhlmann⁵ en 1922. Le cercle et la croix seront ajoutés par Terman dans la formule L du Revised Stanford-Binet de 1937. Le cercle avait déjà servi à Kuhlmann en 1922, il le continua en son test de 1939, mais il laissa tomber le losange.

Les directives, les modifications et le cotage. 1^o Copier le cercle.-- Terman et Amoss s'accordent pour placer cet item à l'âge de trois ans. Dans le C.I.E. l'expérimentateur place la feuille de dessins de la formule d'examen devant le sujet, décrit le cercle avec le doigt, donne un crayon au sujet et indiquant tour à tour le dessin et l'espace en blanc à la droite lui dit: "Dessine cette affaire-là ici." Si le sujet ne réussit pas l'expérimentateur répète tout le procédé et dit: "Maintenant dessine-le encore ici."

Dans le Revised Stanford-Binet l'expérimentateur donne un crayon au sujet et indiquant le cercle dans le livret d'examen il dit: "Fais-en un comme celui-ci. Fais-le ici." L'expérimentateur doit lui faire faire le dessin trois fois en répétant les directives chaque fois. Il ne doit pas permettre au sujet de tracer le modèle.

3 Alfred Binet et Thomas Simon, La mesure du développement de l'intelligence chez les jeunes enfants, Paris, Société Alfred Binet, 1938, pp. 68-69.

4 Terman, The Measurement of Intelligence, op. cit., p. 155 et p. 204.

5 F. Kuhlmann, A Handbook of Mental Tests, Baltimore, Warwick & York Inc., 1922, p. 94, 97, 112.

Amoss semble plus sévère pour le cotage. Pour réussir l'épreuve il faut que la ligne fasse une courbe sur elle-même de façon à renfermer un espace et il ne faut pas que cette ligne fasse plus d'un tour et demie. Terman accorde les points à ce niveau si le sujet fait un mouvement circulaire même si le cercle n'est pas complètement fermé et même s'il est plus elliptique que rond. Il n'oublie pas d'illustrer par des dessins. Pourquoi Amoss n'a-t-il pas fait de même? Le sujet doit réussir une épreuve sur deux dans le C.I.E. contre une sur trois dans le New Revised Stanford-Binet.

Kuhlmann, lui, présente cette épreuve d'une manière différente.⁶ L'expérimentateur place une feuille de papier devant le sujet, fait un ou deux cercles tandis que le sujet regarde, donne un crayon au sujet et continuant à faire des cercles lui dit: "Fais une balle," répétant ces paroles trois fois si nécessaire. Ensuite l'expérimentateur prend la main de l'enfant et lui fait faire quelques cercles. Il libère sa main et répète: "Maintenant fais des balles." L'épreuve est réussie si le sujet fait quelque effort avec assez de succès pour indiquer qu'il essaie de faire un cercle. Ces directives semblent plus élaborées, mais Kuhlmann place cet item à deux ans plutôt qu'à trois.

Au fait, c'est la standardisation établie par un auteur qui donne une réelle valeur à son test; directives et appréciations peuvent varier pourvu qu'elles soient maintenues précises d'une part et comparables à

6 Kuhlmann, A Handbook of Mental Tests, op. cit., p. 94.

un échantillon d'autre part. Terman et Kuhlmann ont surveillé ces facteurs de très près; quant à Amoss, tout ce qu'on a pu trouver de publié sur sa standardisation c'est l'article écrit par le Dr C.E. Stothers à l'appendice 2. Comme nous le verrons dans un chapitre subséquent, cet article renseigne très peu sur la standardisation de l'auteur.

2° Copier une croix.-- Cet item suit l'épreuve du cercle dans le C.I.E. L'expérimentateur ne doit pas tracer la croix avec le doigt mais il doit simplement dire: "Dessine cette affaire-là ici," indiquant l'espace en blanc à côté de la croix. Le sujet réussit cet item s'il dessine deux lignes de n'importe quelle longueur qui se croisent à n'importe quel angle.

Terman introduit cet item à l'âge de trois ans comme test supplémentaire et il le présente d'une façon différente. Amoss se sert de la croix de St-André (+) comme modèle tandis que Terman se sert de la croix de St-Georges (X). L'expérimentateur dessine la croix devant le sujet, après quoi il lui dit: "Fais-en une comme celle-ci." Il n'illustre qu'une fois et le sujet n'a qu'une seule chance.

Contrairement aux découvertes de Gesell, Terman s'est aperçu qu'il était plus difficile de copier la croix que de copier le cercle; 43% des sujets de trois ans ont réussi à copier la croix contre 62% des mêmes sujets qui ont réussi à copier le cercle. La seule condition requise par Terman pour réussir l'épreuve est que l'enfant fasse deux lignes qui se croisent. Aucune attention est accordée aux angles ou à la longueur des

lignes. D'après ses illustrations on voit que l'expérimentateur peut être très indulgent dans le cotage. On peut supposer la même indulgence pour le cotage de cet item dans le C.I.E. bien qu'Amoss n'en dise rien.

30 Copier le carré.— Amoss place cet item à l'âge de quatre ans dans son C.I.E. Les directives sont les mêmes que pour l'item précédent. Il faut qu'un dessin soit réussi sur deux. Voici ce qu'il dit à propos de l'appréciation des réponses: dans un ou l'autre dessin si quatre lignes à peu près droites renferment un espace, on doit accorder les deux mois d'âge mental. Il ne semble nullement préoccupé des angles; d'après ses directives on doit accorder des points même si le carré a plutôt la forme d'un losange ou d'un parallélogramme. Mais alors, dira-t-on peut-être, ce n'est plus un carré!

Dans sa révision de 1916 Terman place cet item au même niveau que dans le C.I.E. et il n'exige qu'une réussite sur trois épreuves. Il nous avertit qu'on ne doit accorder aucune attention à la grandeur du carré car, en général, les jeunes sujets ont une tendance à réduire les figures en les copiant. Il est plus important, souligne-t-il, que les angles droits soient conservés plutôt que les lignes soient droites ou encore que les coins soient entièrement fermés. Les angles droits peuvent être fermés par des lignes qui se croisent quelque peu, mais il ne faut pas que les angles soient arrondis ou qu'ils fassent des 'oreilles' tels qu'illustrés dans ses modèles. Toutefois il ne faudrait pas que la figure soit plus d'une demi-fois plus longue qu'elle est large et les lignes ne doivent pas être brisées bien qu'elles peuvent être courbées.

Remarquons que Terman permet l'usage du crayon plutôt que de la plume. Binet ne permet qu'un essai et il exige la plume. Goddard, selon Terman⁷, permet l'usage du crayon mais il ne permet qu'un essai. Kuhlmann exige la plume et le sujet doit réussir deux dessins sur trois. Cependant ces deux derniers auteurs placent cette épreuve à l'âge de cinq ans. Terman⁸ nous dit que les trois quarts des sujets de quatre ans réussissent une des trois épreuves s'il permet l'usage du crayon et si le cotage est fait avec assez d'intelligence. Il y a une grosse différence dans le résultat si on se sert du crayon ou de la plume et si on exige une ou deux réussites sur trois tentatives.

Cependant Terman a changé d'idée à ce propos car il place ce même item à l'âge de cinq ans dans le New Revised Stanford-Binet. Son expérience ultérieure lui a sans doute démontré que cette épreuve était trop difficile à quatre ans. Il spécifie dans Measuring Intelligence que selon sa méthode de standardisation les items ont été placés au niveau où 50 pour-cent des sujets réussaient.

Il nous semble donc que ce test serait mieux placé dans le C.I.E. à cinq ans plutôt qu'à quatre. Les résultats obtenus par Amoss justifient-ils sa manière de faire? Ou serait-ce parce que Terman avait placé cet item à l'âge de quatre ans qu'il a cru bon de faire de même?

7 Terman, The Measurement of Intelligence, op. cit., p. 156.

8 loco cit.

4° Copier le losange.— Cet item apparaît dans la plupart des échelles d'intelligence. Binet et Simon en font un test de six ans dans la révision de 1911 après en avoir d'abord fait une épreuve de sept ans dans celle de 1908. Ce changement ne semble pas justifié puisque, selon Terman⁹ Binet dit expressément que la moitié seulement des sujets de six ans ont réussi cet item dans les deux révisions de 1908 et de 1911. La plupart des recherches donnent une proportion de succès trop peu élevée à l'âge de six ans pour justifier ce changement surtout si on exige l'usage de la plume. On pourrait placer cette épreuve à six ans à condition qu'on permette l'usage du crayon et qu'on exige une réussite seulement.

Terman place cet item à l'âge de sept ans dans ses deux révisions et il exige deux réussites sur trois essais. Dans la révision de 1916 le sujet doit se servir de la plume tandis que dans celle de 1937 il permet l'usage du crayon.

Kuhlmann¹⁰ emploie cet item de la même façon que Terman dans sa révision de 1916. Il faut que le sujet se serve de la plume et qu'il réussisse deux dessins sur trois.

Amoss aussi place cet item à sept ans dans son C.I.E. Puisqu'il permet l'usage du crayon et qu'il exige une réussite sur deux épreuves il nous semble que cet item pourrait bien être placé à l'âge de six ans.

Ce test aussi bien que celui de copier le carré demande une appréciation des relations d'espace et la capacité de se servir de la perception

9 Terman, op. cit., p. 205.

10 Kuhlmann, op. cit., p. 112.

visuelle pour guider des coordinations motrices très complexes. Cette dernière opération est la plus difficile et elle n'est complètement maîtrisée qu'à l'âge adulte. Il est très intéressant de comparer la difficulté relative de copier un carré et un losange. Les deux ont le même nombre de lignes et les côtés opposés sont parallèles. Mais là où une intelligence de quatre ans est à la hauteur de la tâche pour copier le carré il faut un niveau d'intelligence de sept ans pour copier le losange. Il est probable que personne n'aurait pu soupçonner qu'un changement dans les angles augmenterait à ce point la difficulté de l'item. Il serait très intéressant d'inventer des figures plus compliquées et de les standardiser.

Nous croyons devoir citer ici textuellement ce que Gesell dit dans Mental Growth of the Pre-School Child à propos des dessins à copier. La citation est un peu longue mais elle résume si bien tout ce qu'on peut dire à propos des dessins:

At the age of three years a child is ordinarily able to make the distinction between vertical strokes and horizontal strokes and he is also able to copy a circle from a model. Again he shows no marked interest in this type of circumscribed drawing but if he is held to the task, he shows that he has a certain conception of beginning and end when his circle is complete. He makes a circumscribed curved mark in duplication of a model. Curiously enough, however, we cannot expect him to copy a cross. In another year he is able to copy the cross of St. Andrew, making two straight right-angled strokes, but he has difficulty in completing the oblique cross of St. George.

The ability to copy a square cross and the inability to copy an oblique one presents us with a definite objective cleavage which has considerable psychological interest. From a purely logical point of view these two problems are equally difficult; they require the same amount of energy and both crosses are made up of two short strokes, one applied across

the other. Why is it that one is possible and the other impossible? Is there some incompleteness in the oculo-motor mechanism? Is this incompleteness peripheral or is it central and if it is central, as it possibly is, is the incompleteness due to lack of training or to a lack of native **racially determined neural patterns?**

The copy of a square is somewhat too difficult for the median four-year-old child. The copy of a triangle is a little more difficult. At the age of five years the median child can copy both a square and a triangle but he shows an inability to copy a diamond which recalls a similar lack of mastery over oblique strokes which he displayed a year earlier when called upon to differentiate between the cross of St. George and the cross of St. Andrew 11.

Dans cette série d'items nous trouvons trois épreuves qui semblent ne pas être bien placées dans l'échelle d'Amoss. Copier la croix est trop difficile pour trois ans, copier le carré est aussi trop difficile pour quatre ans et copier le losange trop facile pour sept ans. D'après Gesell l'épreuve copier le triangle à cinq ans serait bien placée dans l'échelle. Nous ne pouvons pas établir de comparaison avec d'autres auteurs pour les items copier le pentagone et l'étoile puisqu'ils ne les ont pas utilisés. Il faudrait donc pour en faire le contrôle, faire passer ces tests à un grand nombre de sujets non-sélectionnés et ensuite établir des normes d'âge.

Cette série d'items que nous venons d'étudier illustre parfaitement le fait que sans une standardisation soignée de procédé et de cotage, le meilleur test peut tromper sur le degré d'intelligence qu'il indique.

11 Arnold Gesell, Mental Growth of the Pre-School Child, New York, The Macmillan Co., 1925, p. 212.

2. La série de pliage de feuilles.

Le pliage de feuilles n'a été employé que par quelques auteurs: Kuhlmann, Terman et Amoss. C'est un test moins populaire que les dessins à copier. Amoss l'introduit dans le C.I.E. aux âges de deux, trois et quatre ans.

L e s s o u r c e s.-- Le pliage de feuilles a été imaginé par Kuhlmann qui l'a employé pour la première fois dans sa révision de l'échelle Binet-Simon en 1922, à six ans et à huit ans. Il se trouve aussi dans sa révision de 1939 à deux, quatre et six ans. Terman l'a employé à l'âge de cinq ans dans la formule L du New Revised Stanford-Binet.

L e s d i r e c t i v e s , l e s m o d i f i c a t i o n s et le c o t a g e.-- En introduisant ce test, Kuhlmann fait le commentaire suivant:

A new test. It is important in such a test that the foldings be of such a nature as not to resemble anything likely to be given as kindergarten training, otherwise the effect of such training would undoubtedly cause individual variations in the results, independently of the stage of mental development ¹².

Kuhlmann¹³ présente ce test en deux parties à l'âge de six ans. Dans la première partie, l'expérimentateur prend une feuille de papier de six pouces carrés et dit: "Je vais plier cette feuille et vois si tu peux plier cette autre (lui montrant une autre feuille) tout à fait comme celle-ci. Regarde bien."

12 Kuhlmann, A Handbook of Mental Tests, op. cit., p. 167.

13 Op. cit., pp. 107-108.

En face du sujet, l'expérimentateur plie la feuille en diagonale pour que le coin de gauche du bas aille recouvrir le coin de droite en haut. Ensuite, il la plie diamétralement pour que le coin de droite en bas aille recouvrir le coin de droite en haut. Enfin, il la plie pour que le coin de gauche en haut aille recouvrir le coin de droite en haut de façon à former un carré de quatre pouces. Après quoi il dit au sujet: "Maintenant, fais-en un comme ça."

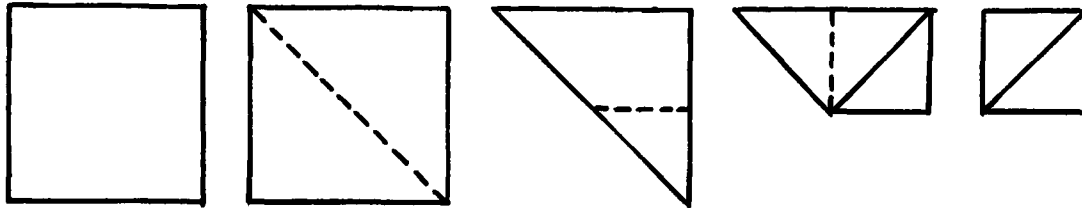


Fig. 2. Les plis à faire dans la première partie du test de pliage de feuilles de Kuhlmann à six ans.

Après que le sujet a plié sa feuille, l'expérimentateur lui dit: "Maintenant, nous allons en essayer un autre." Se servant d'une autre feuille, il plie le coin de gauche en bas sur le coin de droite en haut et replie le coin de droite en haut de la feuille de dessus sur le centre du côté opposé. Enfin, il plie le coin de gauche sur le coin du bas faisant ainsi un triangle à angle droit. Il donne au sujet une autre feuille de papier et lui dit: "Maintenant, fais-en un comme ça."

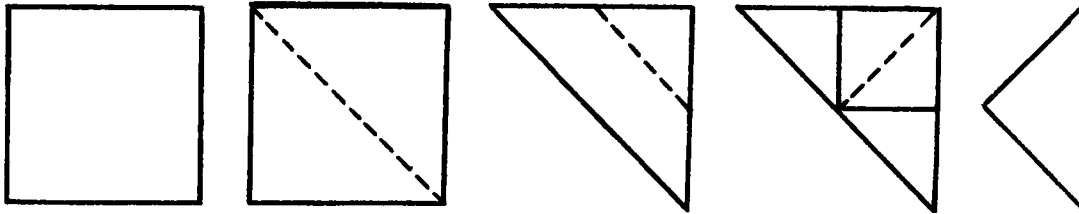


Fig. 3. Les plis à faire dans la seconde partie du test de pliage de feuilles de Kuhlmann à six ans.

L'expérimentateur doit noter le temps que le sujet prend pour plier chaque feuille (le temps maximum étant de 80 secondes) et il doit compter le nombre d'erreurs dans chaque essai. Pour trouver le résultat, il prend la moyenne de temps pour les deux essais (T) et le nombre total d'erreurs (E). Le test est réussi si $T/60 + E = 1.5$ ou moins.

Dans sa nouvelle révision de 1939,¹⁴ Kuhlmann introduit le pliage de feuilles à trois endroits différents: à deux, quatre et six ans. Cette fois-ci on ne tient pas compte du temps pourvu que le sujet plie les feuilles dans le temps maximum de quatre-vingts secondes. Si le sujet réussit à faire quatre des six plis, il passe le test de quatre ans; s'il fait les six, on lui accorde le crédit pour six ans.

Kuhlmann ajoute le pliage de feuilles à deux ans. Il s'agit de plier une feuille en deux, de façon à former un rectangle et de plier une autre fois pour former un carré qui est le quart de la grandeur de la feuille. Si le sujet ne réussit pas la première fois, on répète le

14 F. Kuhlmann, Tests of Mental Developments, Philadelphia, Educational Publishers Inc., 1939, pp. 140, 151-152.

procédé. Le test est réussi si le sujet plie le papier deux fois dans les directions différentes.

Terman¹⁵ n'emploie ce test qu'une fois dans la formule L du New Revised Stanford-Binet à l'âge de cinq ans. Il n'omet pas de faire remarquer, à l'encontre d'Amoss, que ce test est une variante de celui qui a été inventé par Kuhlmann dans son échelle de 1922. Terman exige que la feuille soit pliée diamétralement une fois pour former un triangle et une seconde fois pour former un autre triangle qui sera la moitié moins grand que le premier. Il n'y a donc que deux plis à faire mais l'expérimentateur ne donne qu'une démonstration et le sujet doit réussir après un premier essai. Terman cote ce test assez libéralement. Il n'exige pas que les bords soient parfaitement en ligne et il accorde le crédit pourvu que la forme qui en résulte soit à peu près triangulaire.

Amoss utilise le pliage de feuilles à deux, trois et quatre ans avec cette différence qu'il les fait faire à la suite lorsque le précédent est réussi.

A deux ans, l'expérimentateur donne tout d'abord une démonstration. Il plie une feuille en deux, en donne une autre au sujet et lui demande de faire la même chose. Si le sujet ne réussit pas, l'expérimentateur plie la feuille de ce premier et indiquant les deux feuilles pliées lui dit: "Tu vois, elles sont pareilles."

L'expérimentateur fait disparaître les feuilles et place une autre feuille sur la table. ~~Après~~ avoir attiré l'attention du sujet, il

15 Terman, Measuring Intelligence, op. cit., pp. 92, 217.

plie la feuille en deux comme dans la démonstration et la plie une seconde fois de gauche à droite de façon à former un carré. Plaçant une autre feuille devant le sujet, l'expérimentateur indique du doigt sa feuille pliée et dit: "Fais cela."

Si le sujet échoue, l'expérimentateur recommence le test mais non pas la démonstration. Si dans l'un ou l'autre essai le sujet plie la feuille deux fois de façon à former un carré, on lui accorde le crédit.

Ce test est le même que celui employé par Kuhlmann à l'âge de deux ans. Amoss l'a rendu plus facile en permettant la démonstration d'un pli avant de procéder au test.

A trois ans, il y a un pli d'ajouté au test précédent. La feuille est pliée en quatre et ensuite diamétralement pour former un triangle à angle droit. La technique est la même que celle de deux ans. Si le sujet ne réussit pas la première fois, l'expérimentateur répète tout le procédé une seconde fois. Le test est réussi si le sujet parvient à plier la feuille trois fois pour former approximativement un triangle à angle droit.

Si le sujet réussit cet item, on passe immédiatement au pliage de quatre ans. Cette épreuve est la même que celle employée par Kuhlmann dans la première partie du pliage de feuille de ces deux révisions avec la seule différence qu'Amoss fait répéter la démonstration et l'essai si le sujet ne réussit pas la première fois.

La même question se pose ici comme elle s'est posée pour la série des dessins à copier et comme elle se posera pour les autres qui suivront: jusqu'à quel point la difficulté de ces items est-elle changée par la modification des directives, des procédés et du cotage? Contrairement à

Kuhlmann, Amoss permet une seconde démonstration et un second essai si le sujet échoue la première fois. De plus, jusqu'à quel point l'effet d'apprentissage ou encore la monotonie influencent-ils les résultats du fait qu'on donne tous ces tests à la suite? Ici comme ailleurs, ce n'est pas tant la différence de procédés mais la standardisation sérieuse qui donne la valeur à un test et on ne trouve rien dans le manuel d'Amoss sur son procédé de standardisation.

3. Les images.

Les images ont joui d'une grande popularité et elles ont fait preuve d'une grande utilité dans la mesure du développement mental. Amoss, comme ses prédécesseurs les emploie dans son C.I.E.

L e s s o u r c e s.— Les gravures ont d'abord servi à Binet¹⁶ dans son échelle métrique. Par expérience il a constaté que les réponses obtenues à des questions sur des images sont de trois espèces dont chacune caractérise un niveau intellectuel différent:

1. A trois ans on obtient des réponses par énumération. Ces réponses sont les plus élémentaires. L'enfant énumère isolément les personnages et les objets qu'il reconnaît dans les gravures. Il les énumère sans établir entre eux aucun lien. Il prononce simplement des noms communs. C'est l'objet qui est nommé et non pas l'action. Ce qui caractérise ce type de réponse, c'est qu'il s'agit d'énumération seulement. Le but du test à ce niveau est de se rendre compte si la vue d'un objet familier

16 Binet-Simon, op. cit., pp. 40-46.

dans une gravure provoque la reconnaissance de cet objet. L'enfant normal de trois ou quatre ans est, ainsi que l'appelle Binet, "au stage d'identification", c'est-à-dire qu'il peut identifier les objets familiers dans une image mais non pas les décrire, ne pouvant pas encore saisir les relations qui existent entre eux.

2. A sept ans on obtient des réponses par description. Ici, la nature des personnages et celle des choses sont l'objet d'une désignation; de plus, on fait attention à leurs liaisons, et il en résulte qu'au lieu d'utiliser seulement des mots, le sujet interrogé fait des phrases. La description est supérieure à l'énumération parce qu'elle implique la capacité de réunir les éléments et les qualités d'une image d'une façon assez simple. Ce travail exige un genre plus élevé d'association mentale que la seule énumération. Plus une description est complète, plus elle dénote chez un sujet de la facilité d'association.

3. A quinze ans on obtient des réponses par interprétation. Le sujet de la scène est alors indiqué par un commentaire. Il faut que l'interprétation donnée soit plausible bien qu'il ne faille pas que ce soit nécessairement celle que l'artiste avait dans l'idée.

Les trois gravures employées par Binet pourraient être intitulées: "Le déménagement," "Sans gîte," et "Prisonnier". Binet prétend qu'il suffit que les mots malheureux, déménagement et prisonnier soit insérés dans la réponse pour qu'on ait droit de conclure que le sujet de la gravure a été interprété.

Dans le Stanford-Binet de 1916, Terman¹⁷ se sert des trois images suivantes: "Dutch Home", "River Scene", et "Post-Office." Elles servent pour l'énumération à 3 ans, la description à 7 ans et l'interprétation à 12 ans. Il emploie ces mêmes gravures dans la formule "L" du New Revised Stanford-Binet¹⁸ mais à l'âge de 3 ans seulement. N'ayant pas trouvé ce test satisfaisant pour les deux autres niveaux, il n'emploie pas d'images à 7 ans, et il en choisit une autre "Le Messenger" pour l'interprétation de 12 ans. Dans la formule "M" du New Revised Stanford-Binet, il se sert de trois autres images pour l'énumération et l'interprétation: "Grandmother's Story", "Birthday Party" et "Wash Day" et pour l'interprétation il emploie la même image que dans la formule L.

Les images employées par Binet étaient en couleur; celles de Terman ne sont qu'en noir et blanc. La standardisation de ce dernier a démontré que les couleurs n'étaient pas absolument nécessaires.

Kuhlmann¹⁹ a trouvé que les images dont s'était servi Binet et Terman étaient peu satisfaisantes à l'âge de 3 ans parce qu'elles ne comportaient pas suffisamment de détails familiers aux enfants de cet âge pour se bien prêter à l'énumération. Pour obvier à cette difficulté il se sert d'images dessinées spécialement pour répondre aux besoins de ce test. Il ne les emploie que pour l'énumération et la description.

17 Terman, The Measurement of Intelligence, op. cit., p. 145,190,302.

18 Terman, Measuring Intelligence, op. cit., p. 84.

19 Kuhlmann, A Handbook of Mental Tests, op. cit., p. 157.

Dans sa révision de 1939 Kuhlmann²⁰ n'emploie les images qu'à l'âge de 3 ans.

Dans le C.I.E. Amoss emploie quatre images dont la première pourrait s'intituler "Intérieur canadien". Ce n'est qu'une modification grossière de l'image "Dutch Scene" du Stanford-Binet. Les trois autres images pourraient porter comme titres: "L'escapade" "Cache-cache" et "Le vieux loup de mer".

L a v a l e u r d u t e s t.-- La valeur d'un test comme celui-ci dépend de trois facteurs: 1° le choix des images, 2° la forme des questions (provoquant l'énumération, la description ou l'interprétation), 3° la façon de coter, et 4° la détermination de ce qui constitue une bonne réponse, surtout pour l'interprétation.

Il est évident que le choix des images joue un rôle d'importance primordiale. Certaines images plus que d'autres porteront le sujet à donner un genre de réponse particulier. Les unes se prêtent mieux à l'énumération, d'autres à la description et enfin d'autres à l'interprétation. Il y a très peu d'action et de détail dans les images employées par Binet. Elles sont donc difficiles à décrire. Par contre, les images Jingleman-Jack qu'emploie Kuhlmann représentent des situations si familières et elles possèdent tellement d'action qu'il est rare qu'un enfant normal de 5 ou 6 ans manque de les décrire. Les images de Terman dans le Stanford-Binet ont été choisies parmi un grand nombre parce qu'elles

20 Kuhlmann, Tests of Mental Development, op. cit., p. 142.

se prêtaient bien à l'énumération, à la description et à l'interprétation. Toutefois, il fait remarquer que la deuxième, "The River Scene" se prête plus que les autres à la description.

Tous les auteurs s'accordent pour placer l'énumération à trois ans. Il n'en est pas de même pour la description et l'interprétation. Binet a placé la description à sept ans et on l'a gardée à cet âge dans toutes les révisions excepté celle de Bobertag. Toutefois, les statistiques de différents chercheurs montrent beaucoup de désaccord. Ce manque d'accord est dû au fait que différents investigateurs se sont servis de différentes séries d'images qui n'avaient probablement pas le même degré de difficulté. Comme il a été mentionné plus haut, le genre d'images détermine plus que tout le reste le niveau d'âge auquel on peut obtenir la description ou l'interprétation.

Différents auteurs ont situé l'interprétation de douze à quinze ans. Ici encore, tout dépend du genre d'images, de la façon dont on pose la question et de la manière de coter. Les images Jingleman-Jack qu'emploie Kuhlmann sont aussi faciles à interpréter à dix ans que les images Stanford à douze ans.

La forme de la question influencera aussi la réponse. L'interprétation spontanée ("De quoi est cette image?" ou "Que voyez-vous dans cette image?") ne vient pas plus facilement à quatorze ans que l'interprétation provoquée ("Expliquez cette image") à douze ans.

La façon de coter n'est pas moins importante. Si, avec les images Stanford, par exemple, on exige trois réponses satisfaisantes sur quatre, ce test appartiendra au niveau de douze ans, mais si on en exige que deux

sur quatre, il pourra être placé un ou deux ans plus bas.

Il reste comme dernier point à décider ce qui constitue une bonne interprétation. Il n'y a pas de moment déterminé dans le développement mental où on obtient du coup cette capacité d'interpréter les images. Comme pour le développement de toutes les autres capacités intellectuelles, celle-ci se développe graduellement. Il faut donc décider si telle ou telle réponse contient l'interprétation qu'on a le droit d'attendre au niveau d'âge où le test est placé. Il est donc de prime importance que l'auteur spécifie bien le genre de réponse acceptable au niveau où il a placé le test.

Voyons maintenant comment Amoss a procédé dans cette série d'items. Il s'accorde avec la plupart des auteurs pour placer l'énumération à trois ans, la description à sept ans et l'interprétation à douze ans.

Nous avons vu plus haut jusqu'à quel point le choix d'images peut influencer la valeur du test. Ce choix ne peut être justifié que par une standardisation rigoureuse de l'auteur. Mais ici, comment vérifier puisqu'Amoss n'offre aucune donnée? Il faudrait donc pour faire ce contrôle, prendre sa série d'images et la faire passer à un nombre assez considérable de sujets de trois, sept et douze ans en employant les directives et le cotage indiqués dans le C.I.E. Ensuite il faudrait présenter à ces mêmes sujets la série d'images du New Revised Stanford-Binet en suivant rigoureusement les directives et le cotage de Terman. On pourrait alors comparer les résultats obtenus et établir la validité des items d'Amoss en se servant de ceux de Terman comme critère.

Les directives, les modifications et le cotage.-- Amoss présente tout d'abord la série d'images à trois ans pour en obtenir une énumération. L'expérimentateur présente la première image au sujet et lui dit: "Montre-moi la petite fille." Si le sujet ne répond pas, l'expérimentateur dit en montrant la petite fille: "Tiens, voilà la petite fille." Ensuite il dit: "Maintenant, montre-moi tout ce que tu vois dans l'image et dis-moi ce que c'est." Si le sujet ne réussit pas à énumérer trois objets en une minute, on passe à la deuxième et à la troisième image, en disant chaque fois: "Montre-moi tout ce que tu vois dans cette image et dis-moi ce que c'est." L'item est réussi si dans une des trois images, le sujet a énuméré spontanément trois objets dans l'espace d'une minute. Cette façon de procéder ressemble de très près à celle que Terman a employée dans sa première révision. Toutefois, Terman modifie grandement ce test dans le New Revised Stanford-Binet. L'expérimentateur présente les trois images dans l'ordre indiqué et il dit: "Regarde cette image et dis-moi tout ce que tu peux." Si le sujet ne répond pas on répète seulement: "Dis-moi tout ce que tu peux." Si le sujet arrête après avoir nommé un ou deux objets, on l'encourage en disant: "Dis-moi tout ce que tu peux." Il n'est permis à personne d'encourager le sujet plus d'une fois. Le test est réussi si le sujet énumère spontanément trois objets dans deux des trois images.

Il est évident que Terman est beaucoup plus sévère qu'Amoss pour les directives et pour le cotage. Il l'est avec raison puisque l'expérience lui avait montré que sa première révision était trop facile aux bas-âges. Nous pouvons donc conclure que cet item tel que proposé par

Amoss est trop facile pour trois ans. La seule chose qui pourrait augmenter la difficulté c'est d'avoir déterminé une limite de temps. Amoss exige que cette énumération se fasse en une minute.

Les directives d'Amoss pour la description ressemble de très près à celles des deux révisions de Terman, mais la façon de coter diffère. Terman exige deux réussites sur trois tandis qu'Amoss n'en exige qu'une, mais la description doit se faire en une minute. Terman donne plusieurs modèles de réponses acceptables dans ses deux révisions tandis qu'Amoss nous dit que si le sujet donne une interprétation on lui accorde le crédit ou encore si le sujet donne une description évaluée ainsi: 15% pour (i) chaque adjectif autre que les articles et les démonstratifs, (ii) chaque nom (mère, fille, etc.), (iii) chaque adverbe autre que "là," (iv) chaque verbe intransitif, participe ou infinitif; 20% pour (i) chaque complément du sujet ou du verbe, (ii) verbe et complément; 50% pour une demi-explication. Cette façon de coter est bien onéreuse, bien complexe et manque de clarté. Encore ici il fixe une limite de temps; la description doit se faire en une minute. Il est le seul auteur à tenir compte du temps dans ce genre d'épreuve.

Dans le C.I.E., pour provoquer l'interprétation à douze ans, l'expérimentateur présente les quatre images dans l'ordre indiqué et dit: "De quoi est cette image? Explique l'image." Cette dernière phrase indique clairement au sujet ce qu'on veut de lui. C'est pour échapper à cette difficulté que Terman pose la même question aux trois niveaux différents dans le New Revised Stanford-Binet: "Regarde cette image et dis-moi tout ce que tu peux à propos de cette image."

Terman est sévère pour coter cet item. Il n'y a qu'une image (ceci rend donc l'épreuve beaucoup plus difficile), et pour obtenir le crédit le sujet doit donner les trois idées fondamentales. Il illustre avec des exemples de bonnes et de mauvaises réponses. Amoss accorde un crédit de trois mois pour quatre interprétations, un crédit de deux mois pour trois interprétations et un crédit d'un mois pour deux interprétations. Il se trouve à diminuer beaucoup la difficulté de cet item et le place ainsi à trois niveaux différents. Il est le seul auteur à accorder des crédits partiels à des tests.

De plus, Amoss n'indique pas ce qui peut être considéré comme une interprétation suffisante. Tout est laissé à l'expérimentateur. Comme c'est un des tests où il peut entrer beaucoup de subjectif il y a encore ici une autre raison de douter de sa valeur car il ne peut y avoir d'uniformité dans le cotage.

Ce genre d'épreuves a une grande valeur discriminative et c'est pourquoi il a été conservé malgré les nombreux facteurs qui peuvent influencer les résultats. Les auteurs sérieux en ont assuré la validité par une standardisation sérieuse des directives et des règles de cotage. Amoss n'a-t-il pas vu ces difficultés ou les a-t-il ignorées par pure négligence?

4. Les répétitions de chiffres à la suite.

Un des plus anciens tests d'intelligence pour mesurer la mémoire auditive est sans contredit la répétition de chiffres. C'est un test qui se donne et qui se cote facilement, qui est spécifique quand au genre de capacité qu'il mesure. En dernier lieu, il se prête bien à la standardisation quantitative.

Ce qui distingue la répétition de chiffres de la répétition de phrases c'est son rapprochement marqué à la répétition de syllabes de non-sens, méthode classique d'étudier la mémoire. A cause de leur manque de rapport entre eux, les chiffres ont été utilisés presque universellement pour mesurer l'étendue de la mémoire. C'est pour cette raison que la répétition de trois chiffres taxent autant la mémoire que la répétition de six syllabes dans une phrase.

L e s s o u r c e s.-- Binet emploie la répétition de chiffres à cinq niveaux différents. Elle fera partie de toutes les révisions de l'échelle métrique bien que les auteurs ne s'accordent pas à placer les différentes séries aux mêmes âges.

L a v a l e u r d u t e s t.-- On a souvent mis en doute la valeur de ce test sous prétexte (1) que ce n'est pas un test de pure mémoire mais qu'il dépend en grande partie de l'attention, et (2) que les résultats sont trop influencés par le genre de représentation (type of imagery) de l'enfant. Il est vrai qu'il y a plus d'une fonction mentale qui entrent en jeu dans ce test. On pourrait en dire autant de

tous les tests de l'échelle métrique et même de tous les tests qu'on pourrait inventer. Il est impossible d'isoler une fonction quelconque pour la mesurer séparément. En fait, les diverses fonctions qu'on appelle mémoire, attention, perception, jugement, etc., n'opèrent jamais seules. Il n'y a pas de facultés spéciales correspondant à ces termes qui sont tout simplement des noms commodes qui caractérisent des procédés mentaux de divers genres. Dans n'importe quel test c'est "l'habileté générale", qui parfois agit en grande partie par la mémoire, d'autres fois par le raisonnement ou encore par la discrimination sensorielle, etc.

On alléguera encore que le test est rendu en grande partie invalide par l'existence de différents types de représentations (imagery types). La chose n'est pas prouvée par les faits. Les expériences ont démontré que des types de représentations pures sont très rares et que les enfants surtout sont caractérisés par des représentations mixtes (mixed imagery). Il y a très peu d'enfants qui manquent d'imagination auditive au point où ils seraient placés à un désavantage sérieux dans ce test.

Comme test d'intelligence générale, Wechsler²¹ prétend que d'après ses propres observations c'est un des moins bons. Il l'a quand même conservé dans son échelle pour deux raisons: (1) Bien qu'en général il soit une mesure médiocre d'intelligence, c'est une très bonne mesure aux niveaux inférieurs. Excepté dans les cas de défauts spéciaux ou de

21 David Wechsler, The Measurement of Adult Intelligence, Baltimore, The Williams & Wilkins Company, 1941, p. 85.

maladies organiques il trouve neuf fois sur dix qu'un adulte qui ne peut retenir cinq chiffres à la suite et trois à reculons est un débile mental. (2) Une difficulté marquée à répéter les chiffres à la suite ou à reculons acquiert souvent une signification diagnostique. Le test tirerait donc sa valeur de son pouvoir à distinguer normaux d'anormaux et celui de signaler certaines maladies organiques. Selon lui, la capacité de répéter des chiffres serait une des nombreuses capacités nécessaires au fonctionnement intellectuel mais seulement un minimum est requis. Très développée elle n'indiquerait qu'une bonne mémoire de routine mais non pas une intelligence supérieure car il n'y a qu'une très faible corrélation entre cette habileté à répéter des chiffres et les niveaux supérieurs d'intelligence.

L e s d i r e c t i v e s , l e s m o d i f i c a t i o n s
e t l e c o t a g e.-- Les modifications dans le placement des items aux différents âges, dans les directives et le cotage changeront beaucoup la difficulté des items de ce genre-ci.

La difficulté d'une série de chiffres est augmentée d'une façon assez considérable en y ajoutant qu'un seul chiffre. D'ordinaire, un enfant de quatre ans peut répéter quatre chiffres mais il faudra qu'un enfant soit âgé de sept ans pour répéter cinq chiffres, et qu'il soit âgé de dix ans pour en répéter six.

Il est toujours intéressant de noter les erreurs qui se font dans les répétitions de chiffres. L'erreur la plus commune est d'omettre un chiffre au commencement de la série. Si la capacité de l'enfant est

décidément inférieure au test, il ne donnera probablement que les deux ou trois derniers chiffres de la série. Les substitutions de chiffres sont aussi très fréquentes et si elles forment une série qui ressemble peu à l'original, c'est un mauvais signe qui indique un manque de sens critique que l'on trouve fréquemment dans les personnes d'intelligence inférieure.

Les séries de chiffres de même longueur ne sont pas toujours d'égale difficulté et c'est pourquoi il ne faut jamais improviser des séries de chiffres au moment de l'expérience. Il faut éviter tout particulièrement les séries en ordre ascendant ou descendant, la répétition d'un même chiffre à intervalles réguliers ou encore d'autres particularités qui favoriseraient le groupement des chiffres rendant ainsi la répétition plus facile.

Les directives sont à peu près les mêmes pour tous les auteurs. L'expérimentateur dit au sujet: "Ecoute bien. Je vais te dire des chiffres et quand j'aurai fini, je veux que tu les dises comme moi." L'expérimentateur répète la première série et si nécessaire la deuxième et la troisième. Pour réussir et faut que le sujet répète une série sans erreur.

Les auteurs diffèrent d'opinion sur la vitesse à laquelle on doit répéter les chiffres. Binet coupe les chiffres d'un intervalle d'environ une demi-seconde. Terman trouve que deux chiffres à la seconde est trop rapide. Dans sa révision de 1916, il fait répéter les chiffres un tout petit peu plus vite qu'un à la seconde et dans celle de 1937 il exige qu'ils soient répétés un à la seconde. Kuhlmann emploie aussi la répétition de chiffres. A deux ans les deux chiffres doivent être lus à

une vitesse d'un à la seconde, à quatre ans les trois chiffres doivent être lus en deux secondes et à sept ans les cinq chiffres doivent être lus en trois secondes. Voici ce qu'il dit à propos de la répétition de chiffres:

A Binet-Simon 1908 test. The authors caution the examiner about reading the numerals in the several tests of this sort at the given uniform rate of two numerals a second independently of the length of the series. In the present tests this rate is increased with the length of the series, which, we believe, gives a more favorable condition for correct recall, or at least offers a more uniformly favorable condition for recall in series of different lengths ²².

Il recommandera de lire les chiffres un peu plus vite qu'un à la seconde dans sa révision de 1939.

Des trois méthodes préconisées par Binet, Terman, et Kuhlmann, laquelle Amoss choisira-t-il? Il fait un compromis. Il adopte la méthode de Kuhlmann en variant quelque peu pour ne pas être accusé de l'avoir copié directement et il se sert des séries de chiffres de Terman tout en ayant soin de changer un chiffre par ici, par là. Pour la répétition de trois chiffres à trois ans, Amoss dit que la première et la troisième séries doivent être répétées à une vitesse de trois chiffres en deux secondes et la deuxième série à une vitesse d'un chiffre à la seconde. A quatre ans, même procédé que pour trois ans (Kuhlmann n'a pas de répétition de chiffre à cet âge-ci, alors Amoss fait répéter quatre chiffres à une vitesse de trois chiffres par deux secondes - ce n'est pas chose facile - le quatrième chiffre est fort embarrassant!); à sept ans, cinq

22 Kuhlmann, A Handbook of Mental Tests, op. cit., p. 159.

chiffres en trois secondes; à dix ans six chiffres en quatre secondes, et à quatorze ans sept chiffres en quatre secondes et demie. Pour chaque âge la deuxième série doit toujours être répétée à une vitesse d'un chiffre à la seconde.

Quel motif peut bien guider Amoss pour lui faire changer de procédé à la deuxième répétition. S'il croyait vraiment qu'il lui fallait tenir compte de l'apprentissage il semble vraisemblable qu'il exigerait de ralentir davantage à la troisième répétition pour en augmenter la difficulté contrairement à ce qu'il fait d'en revenir à la méthode de la première série. Il est peu probable qu'il s'en soit préoccupé, et il y a tout lieu de croire qu'il introduit cette petite variante pour ne pas être accusé d'adopter globalement la méthode d'un autre auteur.

En nous référant au Tableau IV nous constatons que les quatre auteurs ne s'accordent pas sur le nombre de chiffres à répéter aux différents âges. Binet et Kuhlmann placent les répétitions de deux et de trois chiffres à trois et à quatre ans, tandis que Terman et Amoss les placent à deux et à trois ans. Tous s'entendent à placer cinq chiffres à sept ans excepté Binet qui les met à huit ans. Goddard suit l'exemple de Binet, mais Kuhlmann les garde à sept ans. Les données de plus d'une douzaine de recherches importantes en Amérique, en Angleterre et en Allemagne s'accordent à montrer que ce test doit rester à sept ans.²³ Amoss ne s'est donc pas trompé en suivant de près la révision de 1916.

23 Terman, The Measurement of Intelligence, op. cit., p. 196.

Amoss imite encore Terman dans le choix même des chiffres. Des sept séries de chiffres qu'emploie Terman à sept, dix et quatorze ans, Amoss en prend trois qu'il garde tels quels et quatre dont il ne change qu'un chiffre (voir le tableau IV). Les séries de chiffres d'Amoss à sept, dix et quatorze ans ne sont donc qu'un mélange à la Terman et à la Kuhlmann; de Terman il emprunte les chiffres, et de Kuhlmann la façon de procéder. Il a sans doute cru obtenir ainsi un test supérieur à celui de ces deux auteurs et qu'il n'avait pas besoin de mettre à l'épreuve. S'il l'a vraiment standardisé, et nous n'en savons rien, et qu'il n'a pas changé les chiffres, c'est donc qu'il manque complètement d'imagination.

Que valent les deux séries de chiffres qu'il propose à trois et quatre ans? A trois ans il n'emploie que les chiffres 1, 2 et 3 qu'il intervertit, et à quatre ans les chiffres de un à cinq seulement. Il est évident que l'emploi d'un nombre restreint de chiffres diminue beaucoup la difficulté des séries. De plus, il les fait répéter en raison de trois en deux minutes pour les différentes séries. Ces deux tests sont donc beaucoup plus faciles que ceux de Terman dans ses révisions de 1916 et de 1937, et par conséquent trop faciles pour les âges auxquels il les a employés.

Amoss change encore d'une autre façon son procédé. Si, par exemple, le sujet réussit la répétition de trois ans on passe immédiatement à celle de quatre ans; si celle-là est réussie on passe à celle de cinq ans et ainsi de suite. Terman met en garde contre ce procédé. Il n'est pas nécessaire, dit-il, et souvent même, il n'est pas opportun de

TABLEAU IV.-

Les séries de chiffres répétés à la suite employées par divers auteurs aux différents âges

Niveau	Echelle métrique de Binet	Terman Révision 1916	Terman - Révision 1937 Formule L	Terman - Révision 1937 Formule M	Kuhlmann Révision 1922	Amoss C.I.E.
2 ans			4-7 6-3 5-8	5-8 7-2 3-9		
3 ans	3-7 6-4 5-9	6-4-1 3-5-2 8-3-7	6-4-1 3-5-2 8-3-7	7-4-9 9-6-1 2-5-3	6-4 8-5 3-7	3-1-2 2-3-1 2-1-3
4 ans	3-0-8 7-1-5 9-2-6	4-7-3-9 2-8-5-4 7-2-6-1	4-7-2-9 3-8-5-2 7-2-6-1	4-7-3-1 5-2-8-3 9-4-1-7	2-6-4 7-5-3 8-1-9	4-1-3-5 2-5-1-4 4-2-5-3
7 ans		3-1-7-5-9 4-2-3-8-5 9-8-1-7-3	3-1-8-5-9 4-8-3-7-2 9-6-1-8-3		6-5-2-8-1 4-9-3-7-5 2-8-6-1-9	3-1-7-5-9 4-2-8-3-5 9-6-8-1-7
8 ans	8-0-3-5-1 7-6-2-9-4 3-5-2-1-7					
10 ans	1-5-3-0-9-4 8-7-3-5-0-6 4-9-2-7-6-1	3-7-4-8-5-9 5-2-1-7-4-6	4-7-3-8-5-9 5-2-9-7-4-6 7-2-8-3-9-4	2-9-4-8-1-6 9-6-2-7-3-8 5-1-7-2-6-9		3-7-4-8-5-9 5-2-1-7-4-6 8-3-7-5-9-2
14 ans	7-3-0-6-2-9-8 4-3-5-1-9-2-7 3-5-9-2-6-7-1	2-1-8-3-4-3-9 9-7-2-8-4-7-5				2-1-8-3-4-6-9 9-6-2-8-4-7-5
Adulte sup.		7-2-5-3-4-8-9-6 4-9-8-5-3-7-6-2 8-3-7-9-5-4-8-2				

LES SÉRIES D'ITEMS

N.B. Les séries d'adulte supérieur II et d'adulte supérieur III des formules L et M de la révision de 1937 ont été omises dans ce tableau.

donner les répétitions de chiffres des différents âges à la suite, sans y intercaler d'autres tests. Bien que ce soit permis avec des enfants plus vieux, le pouvoir d'attention soutenue est si faible chez les jeunes enfants qu'aucun test ne devrait durer plus de deux ou trois minutes. Pour les enfants en bas de six ou sept ans, les tests devront donc être donnés dans l'ordre indiqué au livret de document.²⁴

Il est donc admissible de conclure de cette étude que toutes les séries de chiffres d'Amoss ne sont pas valides si celles de Terman servent de critère. A Amoss de prouver qu'elles le sont.

5. La répétition de chiffres à reculons.

La répétition de chiffres à reculons est une variante de la répétition de chiffres à la suite. Bobertag suggéra l'emploi de ce test mais il ne semble pas qu'il ait été standardisé avant la recherche de Stanford.

L e s s o u r c e s.-- Ce test n'apparaît pas dans l'échelle métrique de Binet. Terman l'emploie dans ses deux révisions ainsi que Kuhlmann. Wechsler l'utilise aussi dans The Measurement of Adult Intelligence.²⁵

L a v a l e u r d u t e s t.-- D'après Terman²⁶ la répétition de chiffres à reculons a une plus grande valeur comme test d'intelligence que la répétition de chiffres à la suite. Elle est moins mécanique et

24 Loc., cit.

25 Wechsler, op. cit., p. 171.

26 Terman, The Measurement of Intelligence, op. cit., p. 208.

exige plus d'attention. Les chiffres doivent être si bien fixés dans la mémoire qu'ils pourront être retenus assez longtemps pour être répétés un à un à reculons.

Ce test est beaucoup plus difficile que la répétition de chiffres à la suite. Il est possible de répéter cinq chiffres à la suite à sept ans et six à dix ans, mais si on renverse l'ordre on ne peut que répéter trois chiffres à reculons à sept ans, quatre à dix ans, cinq à douze ans et six à adulte moyen. Même les personnes intelligentes ont de la difficulté à répéter six chiffres à reculons une fois sur trois.

Les enfants arriérés trouvent ce test particulièrement difficile probablement à cause de son élément de nouveauté. Les instituteurs demandent souvent aux enfants d'écrire des chiffres dictés par eux, et même les plus lents y acquièrent une certaine habileté. La répétition de chiffres à reculons exige une certaine facilité d'adaptation à une tâche nouvelle et c'est justement ce que les arriérés sont incapables de faire.

Il arrive parfois que l'enfant donne un ou deux chiffres après quoi il arrête, ayant complètement oublié le reste de la série dans son effort à s'adapter à la tâche nouvelle et relativement difficile de commencer par le dernier chiffre. Dans de pareils cas les arriérés sont enclins à terminer par n'importe quel chiffre qui leur vient à l'idée.

Il serait intéressant de rechercher la relation qui existe entre ce genre de test et le type de représentation (imagery type). Il faudrait qu'une telle étude se fasse avec des adultes habitués à l'introspection. Il semblerait que le succès peut être favorisé par l'habileté à traduire l'impression auditive en une représentation visuelle pour que les chiffres

énumérés puissent se lire comme dans un livre; ce n'est qu'une probabilité. Toutefois, le succès semble dépendre en grande partie de la capacité à se servir de représentations mentales.

Le degré de certitude du sujet sur l'exactitude de sa réponse est toujours moindre que pour la répétition de chiffres à la suite.

Les directives, les modifications et le cota g e.-- D'après le tableau V on peut constater que Terman, Kuhlmann et Amoss s'accordent sur le nombre de chiffres à répéter aux différents âges. Kuhlmann et Terman tombent aussi d'accord sur la vitesse à laquelle on doit répéter les chiffres, c'est-à-dire un à la seconde. Amoss omet d'en parler. C'est un oubli impardonnable car il y aura beaucoup de diversité si ce point est laissé à la fantaisie de l'examineur.

Des douze séries de chiffres qu'utilise Amoss, neuf sont tirées textuellement de la révision 1916 de Terman. Ce bon Dr Amoss ne se gêne pas pour s'emparer des tests de Terman et les faire siens. Kuhlmann fait preuve de plus de délicatesse. Il n'oublie pas de mentionner que sa répétition de trois chiffres à sept ans est prise textuellement de la révision 1916 de Terman. Il dit aussi que sa répétition de quatre chiffres à neuf ans a été employée de la même façon que celle de Terman à l'exception de quelques petits changements dans les chiffres.

Il n'y a donc pas lieu de douter de la validité de cette série d'items (exception faite de la lacune sur la vitesse à laquelle on doit répéter les chiffres) si on reconnaît celle de Terman mais on ne peut tout de même pas en attribuer le crédit à Amoss.

Tableau V.-

Les séries de chiffres répétés à reculs employés par divers auteurs aux différents âges.

	Terman Révision 1916	Terman - Révision 1937 Formule L Formule M		Kuhlmann Révision 1922	Amoss C.I.E.
7 ans	2-8-3 4-2-7 9-5-8		2-9-5 8-1-6 4-7-3	2-8-3 4-1-7 5-9-6	5-1-3 2-8-3 9-5-8
9 ans	6-5-2-8 4-9-3-7 3-6-2-9	8-5-2-6 4-9-3-7 3-6-2-9	5-8-1-3 9-1-7-4 3-7-2-9	6-4-2-8 5-9-3-8 3-6-2-9	6-2-5-8 4-9-3-7 8-6-9-2
12 ans	3-1-8-7-9 6-9-4-8-2 5-2-9-6-1	8-1-3-7-9 6-9-5-8-2 5-2-9-4-1	5-1-3-9-4 9-2-5-1-6 2-6-3-7-1		
Adulte moyen	4-7-1-9-5-2 5-8-3-2-9-4 7-5-2-6-3-8				4-7-1-9-5-2 5-8-3-2-9-4 7-5-2-6-3-8
Adulte sup. I	4-1-6-2-5-9-3 3-8-2-6-4-7-5 9-4-5-2-8-3-7	4-7-1-9-5-2 5-8-3-6-9-4 7-5-2-6-1-8	4-7-1-9-5-2 5-8-3-6-9-4 7-5-2-6-1-8		4-1-6-2-5-9-3 3-8-2-6-4-7-5 9-4-5-2-8-3-7

LES SÉRIES D'ITEMS

6. La répétition de syllabes.

La répétition de syllabes comme la répétition de chiffres est un test qui a pour but de mesurer la mémoire auditive. Il se donne et se cote facilement et se prête bien à la standardisation quantitative. Contrairement à la répétition de chiffres, il faut faire répéter un plus grand nombre de syllabes parce que le rapport se fait plus facilement entre les syllabes.

L e s s o u r c e s.-- On retrouve ce test dans l'échelle métrique de Binet-Simon.²⁷ Il fera partie de toutes les révisions de l'échelle métrique. Terman l'utilise à deux endroits différents dans sa première révision et à cinq endroits dans les formules L et M de sa révision de 1937. Kuhlmann l'emploie à trois ans et à onze ans dans sa révision de 1922 et à trois et quatre ans dans celle de 1939. Amoss l'emploie à trois, quatre et six ans.

L e s d i r e c t i v e s, l e s m o d i f i c a t i o n s et l e c o t a g e.-- On retrouve seize phrases différentes dans l'échelle métrique dont les premières, très courtes, ne sont utilisées que pour amorcer le sujet s'il reste muet ou intimidé. Binet nous dit qu'on ne doit pas employer toutes les phrases mais de les prendre de deux en deux ou de trois en trois quitte à revenir en arrière en cas d'échec. Bref, on s'occupera surtout du but à atteindre: fixer la limite

27 Binet-Simon, op. cit., p. 48.

de répétition correcte d'une phrase après une seule audition. On arrête l'épreuve dès que deux ou trois phrases ont été incorrectement répétées. Selon Binet un enfant peut répéter six syllabes à trois ans, dix à cinq ans et vingt-six à quinze ans.²⁸ Binet offre donc un bon nombre de phrases sans toutefois bien définir le procédé.

Dans son expérience avec ce genre de test, Kuhlmann trouve que la nature et le contenu des phrases sont aussi importants que la longueur. C'est vrai surtout chez les jeunes enfants.

Il est assez difficile de faire la comparaison de ce test car chaque auteur emploie des phrases différentes. Il est évident que les phrases employées n'auront pas toutes le même degré de difficulté car, en plus de la longueur des phrases, l'intérêt suscité chez le sujet par le contenu, la familiarité du sujet avec les mots employés et la façon de coter entreront en ligne de compte.

Amoss suit la révision 1916 de Terman d'assez près aux âges de trois et six ans. Ce dernier n'a pas de répétition de syllabes à quatre ans, mais si nous faisons la comparaison des phrases d'Amoss à quatre ans avec celles de Terman à quatre ans dans sa révision de 1937 nous verrons que ces phrases illustrent bien ce que nous avons dit plus haut à propos de leur difficulté. Voici les phrases proposées par ces deux auteurs à quatre ans:

(Terman) I like to eat ice-cream. (6 syllabes)
My watch has two hands. (5 syllabes)
Give me just one of them. (6 syllabes)

28 Loc. cit., p. 48.

(Amoss) The little boy found his ball under the kitchen table.(14 syllabes)
 We are going to have a good time after school. (12 syllabes)
 When the sun shines the boys and girls play ball.(10 syllabes)

Terman emploie des phrases de cinq à sept syllabes tandis que celles d'Amoss ont de dix à quatorze syllabes. La première phrase de Terman est très facile mais la deuxième et la troisième le sont moins. Terman exige deux réussites sur trois augmentant ainsi la difficulté de cet item.

Amoss a des phrases plus longues mais plus faciles tant qu'au contenu et il est plus indulgent pour les coter. Si une phrase est répétée sans erreur, il accorde un crédit de deux mois et si elle est répétée avec une erreur seulement il accorde un crédit d'un mois. Il se trouve à placer cet item à deux niveaux différents. La difficulté créée par la longueur de la phrase est contrebalancée par l'indulgence dans le cotage.

Jusqu'à quel point chaque item est-il valide à l'âge proposé? Il n'y a que la standardisation qui nous le dira. Amoss les a-t-il standardisés? Il est permis d'en douter.

7. Les définitions.

Le test des définitions est un autre test qui a été universellement employé pour distinguer différents stades de développement.

L e s s o u r c e s.-- Binet en fut l'auteur et il propose des définitions par l'usage à six ans, des définitions supérieures à l'usage à neuf ans et enfin des définitions de mots abstraits à douze ans. Kuhlmann l'emploiera de la même façon mais à des âges différents: cinq, neuf, et onze ans. Terman l'utilise à cinq et huit ans dans sa révision de 1916;

à cinq, onze, douze et quatorze ans dans la formule L du New Revised Stanford-Binet et enfin à quatre, dix, onze, douze, treize ans et adulte moyen dans la formule M de cette même révision. Amoss l'utilise à cinq, neuf et onze ans dans le C.I.E.

L a v a l e u r d u t e s t.-- Ce test a une grande valeur parce qu'il jette beaucoup de lumière sur la maturité des procédés de perception du sujet. Il devient possible de différencier au moins une demi-douzaine de degrés d'excellence dans les définitions selon la maturité intellectuelle du sujet.

A cinq ans l'enfant peut définir en termes d'usage. Les mots choisis sont tout à fait familiers aux enfants de cet âge. Le but du test n'est pas de savoir si l'enfant connaît le sens des mots qu'on lui demande de définir. Pour le jeune enfant, il y a une différence entre savoir un mot et en donner une définition. On cherche à connaître de quelle façon l'enfant perçoit le mot ou plutôt l'objet qu'il représente, s'il pense à la chose en termes d'usage, d'apparence (forme, grosseur, couleur, etc.), de nature, de composition ou de relation. A neuf ans, il peut le définir en termes supérieurs à l'usage, c'est-à-dire en termes de classification ou de description. A onze ans, l'enfant devrait pouvoir donner la définition de mots abstraits. On a trouvé que cette capacité possède un haut degré de relation à la sorte d'intelligence que les échelles métriques mesurent. Il n'est pas étonnant que les processus mentaux très complexes qui entrent en jeu ici soient très significatifs pour mesurer les différences intellectuelles.

Lorsqu'on demande à un enfant de définir un mot abstrait, on essaie de déterminer s'il peut sélectionner les caractéristiques générales qui sont communes à plusieurs idées et s'il peut les abstraire en négligeant les détails qui n'ont pas de rapport à l'idée générale. Il doit surtout exprimer la relation entre des idées. Ceci suppose des procédés de comparaison, d'abstraction et de généralisation. Il n'est pas nécessaire que la définition soit strictement logique ni que le langage soit très choisi, mais il faut que la définition indique que le sens du mot a été saisi.

Les directives, les modifications et le coteage.-- Binet place les définitions par l'usage à six ans, mais les statistiques de Terman et de presque tous les autres investigateurs indiquent qu'elles sont mieux placées à cinq ans.

Kuhlmann, Terman et Amoss procèdent tous de la même façon pour donner le test mais non pas pour le coter. Kuhlmann exige quatre bonnes définitions sur cinq pour réussir le test; Terman en exige quatre sur six dans sa révision de 1916 et deux sur trois dans les formules L et M de sa révision de 1937; et enfin, Amoss accorde un crédit de deux mois pour cinq des six mots bien définis et un crédit d'un mois pour trois des six mots. Il cote de la même façon à neuf ans mais à onze ans il cote cet item à trois niveaux différents: trois mois de crédit pour quatre définitions sur cinq, deux mois pour trois définitions et un mois pour deux. Les mots que le sujet doit définir sont les suivants: argent, rougeole, marteau, chopine et pitié. A l'encontre de tous les auteurs

qui demandent des définitions de mots abstraits à cet âge-ci, Amoss n'en a qu'un. Ces quatre premiers mots sont beaucoup plus faciles à définir que ceux de Terman: pitié, revanche, charité, envie et justice.

Nous pouvons donc conclure que le choix de mots et la façon de coter d'Amoss rendent ce test trop facile à cet âge-ci, le but étant de mesurer la capacité du sujet à définir des mots abstraits.

8. Le test de vocabulaire.

Le but du test de vocabulaire est de déterminer si un sujet connaît la signification des mots et non pas s'il peut en donner une définition logique et complète. C'est pourquoi nous trouvons un test de vocabulaire en plus des tests de définitions dans les échelles d'intelligence.

L e s s o u r c e s.-- Binet n'emploie pas le test de vocabulaire. Le crédit doit en être attribué à Terman qui propose un vocabulaire de cent mots dans sa révision de 1916. Pour réussir aux différents niveaux d'âge les sujets doivent réussir vingt, trente, quarante, cinquante, soixante-cinq, et soixante-quinze mots correspondant aux âges de huit ans, dix ans, douze ans, quatorze ans, adulte moyen et adulte supérieur. Dans sa révision de 1937 il emploie un vocabulaire de quarante-cinq mots. Pour passer le test les sujets doivent réussir cinq, huit, onze, quatorze, seize, vingt, vingt-trois, vingt-six et trente mots aux âges de six, huit, dix, douze, quatorze ans, adulte moyen, adulte supérieur I, adulte supérieur II et adulte supérieur III.

Kuhlmann prend les cinquante premiers mots du test de vocabulaire de Terman qu'il trouve amplement suffisants puisqu'il ne l'emploie qu'à huit ans.

Amoss présente un vocabulaire de cinquante-cinq mots, vocabulaire qu'il a élaboré lui-même.

L a v a l e u r d u t e s t.-- Contrairement à l'opinion générale, l'étendue du vocabulaire d'un individu n'est pas seulement un indice de sa scolarité mais aussi une très bonne mesure de son intelligence générale. Son excellence comme test d'intelligence vient sans doute du fait que le nombre de mots qu'une personne possède est tout de suite une mesure de sa capacité d'apprendre, de son fonds de connaissances verbales et de l'étendue générale de ses idées. La seule objection qu'on pourrait faire au test de vocabulaire comme mesure de l'intelligence générale c'est que le nombre de mots qu'une personne acquiert doit nécessairement être influencé par les occasions favorables qu'elle a de s'instruire et de se cultiver. Il semblerait injuste pour les personnes illettrées. Ce sont ces considérations qui ont fait omettre le test de vocabulaire de la batterie de tests qui formaient l'échelle de Wechsler²⁹. Il a constaté depuis que c'était une erreur. L'expérience lui a montré que les facteurs tels que la scolarité, etc. influencent l'étendue du vocabulaire d'un individu bien moins qu'on pourrait le supposer. Il fut très étonné de constater que des illettrés et même des individus de nationalité étrangère qui n'avaient acquis qu'un bagage assez limité

29 Wechsler, op. cit., p. 101.

d'anglais étaient moins handicapés par le test de vocabulaire que par d'autres qui paraissaient moins linguistiques. Malheureusement, il ne s'en est rendu compte qu'après qu'un nombre considérable de ses sujets avaient été examinés et il a dû mettre ce test comme supplément seulement.

Les statistiques de Terman³⁰ témoignent que dans la plupart des cas le test de vocabulaire donne un quotient intellectuel qui ne s'éloigne pas plus de 10% de celui obtenu sur toute l'échelle. Des centaines d'enfants qu'il a examinés il ne s'en est pas trouvé un seul d'intelligence supérieure qui avait un vocabulaire très inférieur ni un seul d'intelligence inférieure qui possédait un vocabulaire bien au-dessus de la moyenne. Terman trouve une corrélation qui va de 0,65 à 0,91 avec une moyenne de 0,81 entre le test de vocabulaire et l'âge mental obtenu sur toute l'échelle³¹. Wechsler trouve aussi une corrélation très forte entre le test de vocabulaire et l'échelle toute entière ($\eta = 0,85$ ³²).

Il arrive quelquefois qu'un sujet obtient un résultat un peu supérieur ou inférieur à ce qu'on attendrait de lui à en juger d'après son âge mental. La chose arrivera chez des enfants lents qui vivent dans un milieu très cultivé ou encore chez des enfants très intelligents dont le milieu n'a jamais favorisé le développement du langage. Même dans ces cas l'examineur expérimenté ne se laisse pas facilement tromper car

30 Terman, The Measurement of Intelligence, op. cit., p. 230.

31 Terman, Measuring Intelligence, op. cit., p. 302.

32 Wechsler, op. cit., p. 101.

l'enfant lent d'un milieu cultivé montre sa lenteur intellectuelle dans la qualité sinon dans la quantité de ses réponses tandis que l'enfant intelligent de parents illettrés montre son intelligence dans la justesse et l'à-propos de ses réponses.

Certaines études ont été faites en vue de trouver un système de cotation qui prendrait en considération les différences qualitatives du vocabulaire³³. Bien que ces recherches aient démontré que la chose était possible, elle n'était pas pratique parce que les résultats obtenus n'étaient pas proportionnés à la somme de travail qu'exige cette façon de coter. La corrélation entre l'âge mental et le test de vocabulaire augmentait de 0,91 à 0,94 à partir de sept ans à adulte moyen quand on pondérait la valeur des réponses au lieu de les coter tout ou rien.

La standardisation du test de vocabulaire est bien ce qui donnera sa grande valeur au test. Terman nous dit³⁴ qu'il a choisi le dernier mot de chaque sixième colonne d'un dictionnaire qui contenait à peu près 18,000 mots, les mots les plus communs de la langue. Il se base sur l'hypothèse que cent mots choisis selon une règle arbitraire sera un échantillon assez considérable pour être un indice suffisamment constant du vocabulaire d'un sujet. Plusieurs expériences avec cette liste de mots et d'autres choisies de la même manière ont démontré que cette hypothèse était justifiée. Cinq tests de vocabulaire choisis de cette façon ont été donnés à soixante-quinze sujets et les résultats ont montré que la différence moyenne entre deux de ces tests donnés à la même personne

33 Terman, Measuring Intelligence, op. cit., p. 303.

34 Terman, The Measurement of Intelligence, op. cit., p. 229.

était moins de 5%; ce qui veut dire que n'importe quel de ces cinq tests est suffisamment valide pour être utilisé. Il importe peu de savoir si le vocabulaire d'un enfant est de 8,000 mots plutôt que 7,600; ce qui importe c'est de savoir si ce vocabulaire est approximativement 8,000 et non pas 4,000 ou 12,000 ou un autre nombre bien différent.

Pour trouver le vocabulaire d'un sujet Terman propose de multiplier le nombre de mots réussis par 180. (La liste est faite de 100 mots choisis selon une règle, dans un dictionnaire de 18,000 mots.) Ainsi l'enfant qui définit 20 mots correctement possède un vocabulaire de $20 \times 180 = 3,600$ mots; 50 bonnes définitions donneront un vocabulaire de 9,000 mots, etc.

Wechsler³⁵ emploie la même méthode que Terman pour élaborer la liste de mots de son test de vocabulaire. Il choisit cent mots au hasard dans un des Funk and Wagnall's Standard (School) Dictionaries de la façon suivante: Commencant avec une page de nombre impair, il choisit le mot au haut de la colonne de gauche à chaque cinquième page et il continue ainsi jusqu'à ce qu'il ait obtenu cent mots. Ces cent mots ont été ensuite arrangés en ordre de difficulté d'une façon assez rudimentaire et divisés en deux listes de cinquante mots et chacune a été soumise à des groupes d'expérimentation dont le niveau intellectuel était connu. De cette façon les mots qui ne distinguaient pas bien entre différents niveaux d'intelligence étaient rapidement éliminés. Après une plus longue expérimentation les deux listes furent réunies en une seule de soixante mots et elle fut encore réduite à cinquante mots et finalement à sa longueur présente de quarante-deux.

35 Wechsler, op. cit., p. 100-113.

Dans son introduction au C.I.E., Amoss nous dit qu'il a choisi sa liste de mots pour son test de vocabulaire du langage courant des centres ruraux et urbains plutôt que dans un dictionnaire dans le but d'éviter des expressions littéraires ou locales. Il semble bien que son vocabulaire ainsi choisi a beaucoup plus de chances d'avoir de ces expressions locales qu'il veut éviter. Quels sont ces centres où il est allé chercher ses mots? Son test de vocabulaire ne s'appliquerait qu'à ces centres particuliers plutôt qu'à toute la province d'Ontario. Comment a-t-il procédé dans son choix de mots? Quelle a été sa façon de les standardiser? Ce sont autant de questions qui restent sans réponse.

Les directives, les modifications et le cotage.-- La façon de donner ce test est assez simple. On demande au sujet: "Qu'est-ce qu'une orange?" Quelquefois, les jeunes enfants hésitent à répondre, alors on les encourage en disant: "Tu connais ça des oranges. Tu as déjà vu des oranges. Eh bien, dis-moi ce que c'est qu'une orange." On continue la liste de mots jusqu'au point où le sujet manque cinq ou six mots de suite.

Amoss propose cinquante-cinq mots dans son vocabulaire et il établit son système de cotage de la façon suivante: à dix ans il accorde un crédit de deux mois pour vingt bonnes réponses et un crédit d'un mois pour dix-sept; de dix à douze ans un crédit de trois mois pour trente bonnes réponses, deux mois pour vingt-huit et un mois pour vingt-cinq; de douze à quatorze ans, un crédit de quatre mois pour quarante bonnes réponses, trois mois pour trente-huit, deux mois pour trente-six et un mois pour

trente-quatre; de quatorze à dix-sept ans un crédit de six mois pour cinquante bonnes réponses, cinq mois pour quarante-sept, et deux mois pour quarante-quatre. La même question se pose toujours: sur quoi Amoss se base-t-il pour coter ainsi? Il n'en dit qu'un mot dans son introduction. Il prétend qu'en accordant des crédits pour des réponses partielles, il obtient un résultat un résultat gradué plus finement et il élimine la tendance de l'examineur à être trop indulgent dans l'interprétation des réponses. S'il voulait éviter le risque de trop d'indulgence dans l'interprétation des réponses il aurait pu faire comme tous les auteurs qui proposent des modèles de bonnes et de mauvaises réponses.

Dans un travail scientifique, il ne faut rien laisser au hasard si on veut avoir un bon instrument de travail. Amoss ne l'a probablement pas compris puisqu'il néglige ce point capital.

9. Les ressemblances et les différences.

Tous les experts s'accordent à louer la valeur de ce test parce qu'il est tellement peu influencé par l'expérience scolaire.

L e s s o u r c e s.-- Binet présente le test des ressemblances dans son échelle de 1905 à huit ans, mais il n'emploie le test des différences que dans sa révision de 1908 au niveau d'adulte, à la suggestion de Bobertag. C'est à Terman que doit être attribué le crédit de la standardisation de ce test. Il présente le test des différences à sept ans et le test des ressemblances à huit ans et à douze ans dans la ré-

vision de 1916. Il les emploie à sept ans, huit ans, onze ans et adulte moyen dans la formule L et à six ans, huit ans, neuf ans, onze ans, adulte moyen et adulte supérieur II dans la formule M de sa révision de 1937. Kuhlmann ne les utilise qu'à huit ans. Amoss les présente à sept ans, huit ans, douze ans, adulte moyen et adulte supérieur dans le C.I.E.

La valeur de ces tests.-- Le mérite de ces tests vient de ce qu'ils sont faciles à donner et intéressent beaucoup les adultes. Ce sont des tests reconnus par tous les investigateurs comme ayant une grande quantité de "g"³⁶. De plus, ils sont aptes à jeter de la lumière sur les processus mentaux du sujet. La pensée est essentiellement l'association d'idées basée sur des différences et des ressemblances. Presque tous les processus mentaux des plus complexes aux plus simples impliquent jusqu'à un certain degré ces deux genres d'association. Ils sont compris dans les jugements simples des enfants, dans l'appréciation de jeux de mots, dans les inventions mécaniques, dans la classification scientifique de phénomènes naturels et jusque dans l'origine des hypothèses scientifiques ou philosophiques.

La capacité de noter des différences semble précéder quelque peu celle de noter des ressemblances. La difficulté du test vient du fait que les objets à comparer ne sont pas présents aux sens, ce qui veut dire

36 Stern définit l'intelligence comme étant: "A general capacity (G) which may be directed in any field of knowledge". La lettre "g" désigne cette capacité générale qui serait l'intelligence. Un test qui a une forte dose de "g" serait donc une bonne mesure d'intelligence.

que l'enfant doit susciter des idées pour en faire la comparaison ou le contraste. Le sujet peut échouer à cause d'un manque de capacité de représentation, d'une insuffisance d'association ou encore des deux réunies. Il semble probable que les deux facteurs entrent en jeu.

Le développement intellectuel se manifeste surtout par la capacité de noter des différences et des ressemblances essentiels en opposition avec celles qui sont banales, superficielles ou accidentelles. Il y a évidemment beaucoup de différence dans la maturité intellectuelle d'un sujet qui dit qu'une orange et une banane se ressemblent parce que les deux ont des pelures et un autre qui dit que les deux sont des fruits. Ce n'est que vers l'âge d'adulte que la capacité de donner des différences et des ressemblances essentielles est marquante et c'est un très bon signe si un enfant de huit ou neuf ans donne une comparaison de ce genre.

Il est toutefois remarquable de constater le pourcentage d'adultes qui ne dépassent pas le genre superficiel de réponses. C'est pourquoi, contrairement aux méthodes jusqu'ici préconisées, Wechsler³⁷ fait une distinction entre les réponses inférieures et les réponses supérieures en donnant un crédit de deux points à ces dernières et d'un point aux premières.

Wechsler nous dit encore que ce test renseigne d'une autre façon sur le niveau intellectuel. Certains sujets peuvent obtenir un résultat global relativement haut tout en ayant que des crédits d'un point tandis que d'autres sujets ont une proportion plus variée de crédits de zéro,

37 Wechsler, op. cit., p. 88-89.

un et deux. Les premiers ont une bonne capacité intellectuelle mais ils appartiennent à la catégorie de laquelle on ne peut attendre un haut degré de travail intellectuel; les derniers bien inégaux dans leur rendement ont plus de possibilités.

Wechsler obtient une corrélation très élevée entre ce test et les autres de son échelle, et la corrélation de ce test avec l'échelle entière est de 0,73.

Un point reste à noter. Il est bon de remarquer en donnant le test s'il y a tendance à séréotyper, c'est-à-dire s'il y a une tendance à répéter la même idée dans des réponses successives, v.g. "un papillon est plus gros qu'une mouche," "une pierre est plus grosse qu'un oeuf," "le bois est plus gros que la vitre," etc. Dans ce cas il faut coter la réponse comme un échec même s'il arrive que l'idée s'applique à toutes les parties de la question. Règle générale, on trouve ce genre de réponses chez les très jeunes enfants ou chez des enfants plus vieux mais arriérés.

Il serait intéressant de standardiser un nouveau test de ce genre pour les âges plus avancés et qui serait tout spécialement adapté pour différencier la capacité des individus à donner des différences et des ressemblances essentielles.

La technique et l'appréciation des réponses.-- Tous les auteurs s'accordent à placer les différences entre objets à sept ans. La technique est la même pour tous. Binet, Kuhlmann et Terman exigent deux bonnes réponses sur trois pour

réussir et Terman n'en exige que deux sur quatre dans sa nouvelle révision. Amoss exige trois bonnes réponses sur quatre pour accorder deux mois de crédit, et deux sur quatre pour un mois. Il se montre ici plus sévère que ses prédécesseurs. Ce test compense ainsi pour d'autres où il l'est moins. Il reste à noter que deux des quatre items de ce test sont choisis textuellement de la révision 1916 de Terman.

Les ressemblances sont présentées à huit ans par tous les auteurs. Mais là où Terman et Kuhlmann n'exigent que deux bonnes réponses sur quatre pour réussir, Amoss en exige trois sur quatre pour accorder deux mois de crédit et deux sur quatre pour un mois. Ici encore, Amoss est plus sévère.

Terman et Amoss sont les seuls à présenter les ressemblances entre trois choses à douze ans. Terman n'exige que trois bonnes réponses sur cinq pour accorder le crédit tandis qu'Amoss en exige cinq sur cinq pour accorder les trois mois de crédit, quatre sur cinq pour deux mois, et trois sur cinq pour un mois. Deux des cinq items d'Amoss sont tirés textuellement de la révision 1916 de Terman. Le dernier item d'Amoss demande la ressemblance entre trois mots abstraits: bonheur, peine et honte. C'est évidemment plus difficile que de trouver la ressemblance entre des choses concrètes.

Sur ces trois tests de Terman et d'Amoss il y a donc une variante de quatre mois d'âge mental sur la façon de coter seulement. C'est beaucoup. Amoss est-il justifiable de coter aussi sévèrement? Nous serions plus portés à donner raison à Terman puisqu'il a fait une standardisation soignée.

Amoss emploie les différences à trois autres endroits: aux âges de XII - XIV, XIV - XVII et XVII - XX. Binet et Terman les utilisent à adulte moyen. Il faut trouver ici les différences entre des mots abstraits tels que évolution et révolution, paresse et oisiveté, caractère et réputation, pauvreté et misère. Il est évident que le facteur de langage entre pour une large part ici et que le sujet illettré se trouve à un grand désavantage.

Wechsler avait tout d'abord omis le test des ressemblances dans sa première sélection de tests à cause de l'opinion erronée qu'il serait trop influencé par le langage. Ce fut son expérience avec les items de ressemblance d'adulte moyen dans le Stanford-Binet qui le faisait penser ainsi. Parce qu'il avait trouvé un trop grand nombre d'adultes qui échouaient à donner la différence entre "oisiveté" et "paresse", "évolution" et "révolution", parce que les mots "oisiveté" et "évolution" leur étaient inconnus, il doutait de la possibilité d'élaborer un test de ressemblances qui ne serait pas entravé par des difficultés linguistiques. Un peu de recherche lui démontra qu'il y avait possibilité d'augmenter la difficulté des items sans avoir recours à des mots difficiles et trop peu connus. Il demande, par exemple, la ressemblance entre une mouche et un arbre et il trouve que cette question est bien répondue par moins de 20% de la population adulte. Quand de tels mots sont choisis le test de ressemblances peut être complètement libéré des difficultés linguistiques.

Cet exposé de la façon de procéder de Wechsler a pour but de prouver que les trois tests de ressemblances d'Amoss à XII - XIV, XIV - XVII et XVII - XX ans n'ont pas beaucoup de valeur comme tests

d'intelligence parce qu'ils sont trop influencés par le langage. A XII - XIV ans il demande la différence entre péché et crime, argent et richesse, attente et espérance, calme et patience, pauvreté et misère; à XIV - XVII ans, la différence entre paresse et oisiveté, caractère et réputation, probable et possible, avoir et posséder, malheur et accident, noblesse et dignité, émigrant et immigrant; à XVII - XX ans, la différence entre homme d'état et politicien, savoir et sagesse, évolution et révolution, évidence et témoignage, contagion et infection, novice et amateur, hérésie et schisme. Il n'y a que l'érudit qui pourra faire assez bonne figure sur ces tests, tant qu'à l'illettré, il sera évidemment à un très grand désavantage.

Des six items sur les ressemblances et les différences il y en a trois qui sont cotés trop sévèrement et trois qui ont une trop grande difficulté linguistique. Dans une seule série de tests c'est déjà beaucoup d'avoir six tests dont la validité est douteuse.

10. Les problèmes d'arithmétique.

La capacité de résoudre des problèmes d'arithmétique a été reconnue depuis longtemps comme un indice de capacités intellectuelles. Avant l'introduction de la psychométrie on les utilisait comme mesure d'intelligence. Maintenant la plupart des échelles d'intelligence comprennent des items qui demandent du raisonnement mathématique sous une forme quelconque.

L e s s o u r c e s.-- Binet utilise les problèmes d'arithmétique à neuf ans. L'enfant joue au marchand avec l'expérimentateur qui lui achète des boîtes et le sujet doit rendre la monnaie. Terman, Kuhlmann et Amoss reprennent ce test mais dans leur cas les problèmes sont posés oralement. Kuhlmann présente d'autres problèmes à onze ans et Terman en introduit de différents à quatorze ans dans sa révision de 1916 mais ils apparaîtront à adulte moyen dans la formule I de la révision de 1937. Wechsler présente une série de dix problèmes dans son échelle, problèmes qu'on pose au sujet jusqu'au point où il en manque trois de suite. Amoss fait usage des problèmes d'arithmétique à neuf, dix, douze, quatorze ans, adulte moyen et adulte supérieur.

L e u r v a l e u r.-- En plus d'être une bonne mesure d'intelligence générale, les problèmes d'arithmétique ont l'avantage d'être facilement inventés et standardisés. Cependant, leur mérite est diminué parce qu'ils sont influencés par la scolarité, les fluctuations de l'attention et les réactions émotionnelles et passagères du sujet.

A première vue, on serait porté à croire que ce sont simplement des tests de scolarité. Il est vrai que pour les résoudre, le sujet fait appel aux connaissances qui sont ordinairement obtenues à l'école, mais ces connaissances (c'est-à-dire de lecture, d'addition, de soustraction, de multiplication et de division) sont acquises par la plupart des adultes qui ne sont pas débiles et même par plusieurs qui le sont. Le succès dépend cependant de la capacité à appliquer ces connaissances avec facilité et précision aux problèmes donnés - précisément le genre de capacité

qui, si elle manque chez le sujet, ne peut lui être donnée par la scolarité. On peut enseigner même à des arriérés la lecture, l'addition, la soustraction, la multiplication et la division avec un degré passable de précision mais la difficulté vient quand il s'agit de décider laquelle de ces opérations ils doivent choisir pour résoudre le problème. Ils peuvent maîtriser les opérations de routine mais pour ce qui concerne l'initiative, le jugement, et le pouvoir de raisonnement, ils ne sont que très peu éducatibles. La psychologie et la pédagogie de la débilité mentale peuvent être résumées dans cette dernière affirmation. Terman a trouvé que des hommes sans pratiquement aucune scolarité mais d'intelligence à l'adulte moyen passaient ce test aussi fréquemment que des élèves de troisième et de quatrième année d'école supérieure³⁸.

Les directives, les modifications et le cota g e.-- Il a été dit plus haut que Kuhlmann et Terman avaient modifié ce test que Binet proposait dans son échelle métrique à l'âge de neuf ans. Selon leurs exigences les problèmes doivent être donnés oralement et ils doivent être résolus oralement sans l'aide papier ou de crayon. Amoss comme les autres auteurs s'accorde pour placer ce test à neuf ans. Il emploie les mêmes problèmes que Terman avec la seule différence qu'il change quelques mots: au lieu de parler de bonbons, il parle d'acheter des timbres, un livre de lecture et un crayon. Le fond des problèmes reste le même que celui de Terman. Kuhlmann et Terman exigent

38 Terman, The Measurement of Intelligence, op. cit., p. 241.

deux bonnes réponses sur trois pour réussir tandis qu'Amoss exige trois bonnes réponses sur trois pour accorder deux mois de crédit et deux bonnes réponses pour un mois. Il augmente ainsi la difficulté de l'item par sa façon de coter.

A quatorze ans Amoss propose exactement les mêmes problèmes que ceux que présente Terman à quatorze ans dans sa révision de 1916. Il modifie cet item en présentant un problème de plus. Terman placera ces problèmes à adulte moyen dans sa révision de 1937 parce que l'expérience avec sa première révision lui avait montré que les tests à la partie supérieure de son échelle précédente étaient trop difficiles. Terman n'omet pas de dire qu'il a choisi ces problèmes d'une liste utilisée par Bonser dans A Study of the Reasoning Ability of Children in the Fourth, Fifth and Sixth School Grades³⁹.

Terman exige deux réussites sur trois pour accorder les quatre mois de crédit tandis qu'Amoss accorde un mois de crédit pour chacun des quatre problèmes réussis. Des deux tests d'arithmétique de neuf ans et de quatorze ans il y a une différence de trois mois d'âge mental rien que dans la façon de coter les items.

De plus, Amoss emploie des problèmes d'arithmétique à dix ans, douze ans et à adulte moyen. Les problèmes de dix et douze ans n'exigent que des opérations de multiplication et de division tandis que ceux d'adulte moyen sont très complexes. Il emploie toujours son système de crédits partiels pour les coter. Ici encore il faudrait éprouver la

39 Loc. cit., p. 321.

validité de ces items en établissant la corrélation avec d'autres items comme le test de vocabulaire, les ressemblances, etc. ou avec le résultat global du test. Il y a lieu de douter qu'Amoss l'ait fait.

11. Les séries mathématiques.

Les séries mathématiques qu'Amoss présente dans le C.I.E. sont tirées de son Mathematical Ability Test qu'il publia en 1937.

Voici en quels termes Amoss présente son test au revers de la feuille de cotage:

This examination is not a test of mathematical attainment but of mathematical ability and aptitude. . . . This examination is based on the thesis that a test of mathematical intelligence or indeed of any form of intelligence is best obtained by confronting the candidate with a sequence or sequences of situations involving increasingly complex and increasingly disparate relationships of the same kind as determined by inherent indices of gradation and established by usage, rather than with an array of widely selected samplings arranged in order of difficulty only by trial and error standardization and inevitably subject to the influence of local learning.

. . . Results interpreted according to the Table of Mathematical Ages give a general set up corresponding with scores obtained with the Binet-Simon Examination in elementary schools and the Otis Group Intelligence Tests in normal schools. Individual scores with corresponding set ups are, however, not necessarily correlative ⁴⁰.

Amoss choisit les huit séries les plus difficiles des vingt-quatre séries du Mathematical Ability Test et les présente à XIV-XVII ans (adulte moyen et à XVII-XX ans (adulte supérieur). A adulte moyen le sujet reçoit deux mois de crédit pour chaque série réussie ou un mois de crédit pour chacune qui n'a pas plus d'une erreur mécanique. A adulte supérieur, le

⁴⁰ Harry Amoss, Mathematical Ability Test, Toronto, The Ryerson Press, 1937.

sujet reçoit deux mois de crédit pour quatre séries réussies et deux mois pour chaque série additionnelle complètement réussie ou un mois pour chacune partiellement réussie.

Le crédit maximum qu'un sujet pourrait obtenir sur ce test serait donc de seize mois à adulte moyen et dix mois à adulte supérieur. Donner une valeur de vingt-six mois à un test semble fort exagéré. Il est vrai que nous n'avons jamais rencontré un sujet qui ait réussi les huit séries en huit minutes, mais un sujet qui en réussirait cinq obtiendrait un crédit de quatorze mois. Le maximum de crédit qu'accorde Terman pour un test quelconque (à l'exception du test de vocabulaire) est de neuf mois⁴¹. Dans tous les autres tests d'adulte moyen et d'adulte supérieur du C.I.E. le crédit maximum qu'un sujet peut obtenir est de six mois. Alors pourquoi accorder seize mois et dix mois pour celui-ci à ces deux niveaux? Amoss attacherait-il plus d'importance à ce test particulier parce que c'est sa création personnelle?

Il semble qu'Amoss aurait mieux fait de proposer un nombre plus restreint de séries, cinq par exemple, et accorder un crédit de six mois à adulte moyen si trois des cinq séries sont réussies et le même crédit à adulte moyen supérieur pour quatre des cinq séries réussies. Même là, ce serait accorder trop d'importance aux mathématiques puisqu'il y a déjà une série de problèmes de raisonnement mathématique à ces deux niveaux.

41 C'est le test sur les problèmes d'ingéniosité qui sont placés à trois niveaux différents. Si le sujet réussit un problème il reçoit deux mois de crédit à quatorze ans, s'il en réussit deux il reçoit deux autres mois de crédit à adulte moyen et s'il en réussit trois il reçoit cinq mois de crédit à adulte supérieur II.

12. Compter des sous.

Compter des sous est un test qui a été employé à deux niveaux différents: compter quatre sous à quatre ou cinq ans et compter treize sous à six ans.

L'enfant apprend à compter avant de pouvoir appliquer ce procédé à des objets, c'est-à-dire avant de pouvoir compter correctement un groupe d'objets. Compter devient vite mécanique mais compter des objets exige un certain contrôle de ce procédé mécanique.

Ces deux tests apparaissent tout d'abord dans l'échelle métrique de 1908 à cinq ans et à sept ans. Binet place le deuxième test à six ans dans sa révision de 1911. Tous les auteurs s'accorderont pour laisser ce deuxième test à six ans mais ils ne s'accordent pas pour le premier.

Binet et Kuhlmann le place à cinq ans mais Kuhlmann exige que le sujet réponde à la question: "Combien y en a-t-il?" en plus de compter les sous correctement. Terman le place à quatre ans et le test est réussi si le sujet compte les quatre sous correctement une fois. Il nous dit que trois séries de tests préparées pour la révision Stanford ont montré que cet item appartient à quatre ans plutôt qu'à cinq⁴². Toutefois Terman le proposera à cinq ans dans la formule L du New Revised Stanford-Binet mais il en augmentera la difficulté en exigeant que le sujet compte quatre blocs, quatre perles carrées et quatre sous et qu'il réussisse deux fois sur trois.

42 Terman, The Measurement of Intelligence, op. cit., p. 155.

Amoss présente cet item à quatre ans mais le sujet n'a qu'à compter trois sous. Il en diminue ainsi la difficulté car le contrôle du procédé mécanique est beaucoup moins difficile à exercer pour trois objets que pour quatre. Le sujet a moins de chance de se tromper soit en comptant le même sou deux fois ou en passant par-dessus un. Il arrive parfois qu'un sujet obtient la bonne réponse en sautant par-dessus un et en comptant un autre deux fois. Il faudrait compter ce cas comme un échec bien qu'Amoss ne le spécifie pas.

Amoss rend ce test encore plus facile par sa façon de coter. Le sujet réussit s'il compte correctement trois sous une fois en montrant avec le doigt ou deux fois sans montrer avec le doigt. Il peut fort bien arriver qu'un sujet sache compter jusqu'à trois seulement et qu'il le fasse d'une façon tout à fait mécanique sans que pour cela il puisse compter les objets. Selon les directives d'Amoss il faudrait accorder le crédit si le sujet compte jusqu'à trois, deux fois même s'il est évident qu'il n'a pas maîtrisé ce contrôle moteur. Le but de ce test ne serait donc plus de savoir si le sujet peut compter trois objets mais s'il peut compter jusqu'à trois tout simplement.

Les statistiques de tous les auteurs s'accordent pour placer le test: compter treize sous à six ans. Amoss comme les autres le place à six ans mais il en rend la validité douteuse en introduisant une variante dans sa façon de coter. Il accorde une réussite si le sujet compte les sous une fois en indiquant du doigt ou s'il les compte deux fois sans indiquer du doigt. Il se peut fort bien qu'en comptant, le sujet passe un sou et qu'il en compte un autre deux fois. S'il en était ainsi il

faudrait compter le résultat comme un échec mais en proposant cette dernière façon de coter il est impossible de contrôler comment le sujet arrive au résultat.

Voici donc deux autres tests dont il y a lieu de douter la validité tels qu'ils sont proposés par Amoss.

13. Les jours de la semaine et les mois de l'année.

Le test: nommer les jours de la semaine apparaît à neuf ans dans l'échelle métrique de 1908 mais il est omis de la révision de 1911. Kuhlmann l'omet aussi. Goddard le place à huit ans tandis que Terman le place à sept ans dans sa révision de 1916 parce que, dit-il, les statistiques de toutes les recherches américaines justifient sa place à sept ans⁴³. Cependant, il le présentera à huit ans dans la formule L de sa révision de 1937.

Le deuxième test: nommer les mois de l'année apparaît à l'âge de dix ans dans l'échelle métrique de 1908 et à neuf ans dans la révision de 1911. Goddard le place à neuf ans tandis que Kuhlmann et Bobertag l'omettent. Terman le présente comme test supplémentaire à neuf ans dans sa révision de 1916 mais il l'omettra dans celle de 1937.

On a souvent critiqué ces deux tests parce que, disait-on, ils sont en étroite dépendance de la mémoire et passibles d'entraînement. On a prétendu qu'un enfant pouvait passer le premier test sans avoir une conception suffisante de "semaine", "hier", ou "avant-hier". Cette critique

⁴³ Terman, The Measurement of Intelligence, op. cit., p. 201.

est justifiée si le test est donné selon l'ancienne méthode, mais elle ne s'applique pas si on le donne à la façon que Terman le propose, c'est-à-dire avec deux vérifications: "Quel jour vient avant mardi?" et "Quel jour vient avant vendredi?" Ces deux questions permettent de distinguer les réponses données de mémoire de celles qui sont basées sur une connaissance véritable du temps.

Les esprits lents et les débiles mentaux ne font pas bonne figure dans ce premier test parce que leurs associations mentales sont plus faibles et moins nombreuses que celles de personnes normales. Il en résulte que les expériences dont ils se souviennent sont moins bien localisées dans le temps relativement avec d'autres événements.

Il semblerait que l'âge indépendamment de l'intelligence a un certain effet sur l'habileté à nommer les mois de l'année. Les adultes d'un niveau intellectuel de neuf ans réussiraient ce test aussi bien que des enfants normaux de dix ans⁴⁴. C'est sans doute la raison pour laquelle Terman l'omettra dans sa révision de 1937.

Amoss, comme Terman, présente le premier test comme test supplémentaire à sept ans. Il suit textuellement Terman dans sa présentation mais il diffère dans sa façon de coter. Terman n'accorde aucun crédit si les jours seulement sont répétés sans les vérifications tandis qu'Amoss accorde un mois de crédit pour la répétition des jours et un mois pour deux vérifications. Il semble que donner ainsi, le test est bien placé à sept ans car Terman constatera qu'il est trop difficile pour sept ans

44 Terman, op. cit., p. 252.

et il le placera à huit ans dans sa révision de 1937.

Amoss place le deuxième test à neuf ans comme le fait d'ailleurs Terman. Il ne le présente pas comme test supplémentaire mais comme test régulier. Tandis que Terman exige que les mois soient nommés dans l'espace de quinze à vingt secondes avec au plus une erreur et deux sur trois vérifications réussies, Amoss ne parle pas de limite de temps mais il faut que le sujet nomme les mois sans erreur pour obtenir un mois de crédit et qu'il réussisse deux vérifications sur trois pour obtenir le second mois de crédit.

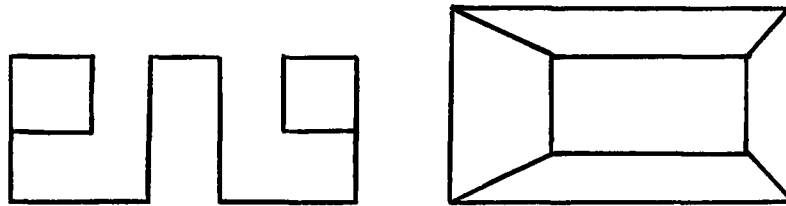
C'est toujours la façon de coter d'Amoss avec ses crédits partiels qui vient changer la difficulté des items et les résultats du test. Pour les six tests d'une année qui seraient partiellement manqués, un sujet pourrait obtenir six mois de crédit sur le C.I.E. là où il n'obtiendrait rien sur le Stanford-Binet.

14. Reproduire des dessins de mémoire.

Reproduire des dessins de mémoire est un autre test qui a été utilisé fréquemment dans la mesure de l'habileté mentale.

L e s s o u r c e s.-- Binet utilisa deux dessins à reproduire de mémoire dans son échelle de 1905, une section de prisme et une grecque, mais il les supprima dans celle de 1908 pour ensuite les présenter de nouveau dans sa révision de 1911 à dix ans. Terman présente les mêmes dessins à dix ans et Kuhlmann fait de même dans sa révision de 1922. A l'exception de Goddard qui recommande que ce test soit placé à onze ans, tous les

auteurs semblent s'accorder pour le placer à dix ans. Selon les données de Terman, le test sera bien placé à dix ans ou à onze, le tout dépendant de la façon de coter. Amoss placera ce test à dix ans et il ajoutera de plus deux autres dessins à reproduire à huit ans et à adulte moyen.



La grecque

La section de prisme

Fig. 4. Les dessins à reproduire de mémoire à dix ans.

L e u r v a l e u r.-- Binet nous dit que les principaux facteurs de succès dans ce genre de test sont: l'attention, la mémoire visuelle et un peu d'analyse. Selon Terman⁴⁵ la capacité d'analyse rapide semblerait le facteur le plus important parce que, si les dessins sont analysés, ils pourront être reproduits par la mémoire verbale de l'analyse. En effet, sans une analyse quelconque, il serait pratiquement impossible de se souvenir des dessins puisque l'un d'eux contient treize lignes et l'autre, douze. Le succès n'est possible qu'à condition de grouper les lignes selon leurs relations pour que plusieurs d'entre elles soient perçues comme une unité. De cette façon, la section de prisme composée de douze lignes peut être réduite à quatre éléments: (1) le rectangle extérieur, (2) le rectangle intérieur, (3) la position

45 Op. cit., p. 260.

du rectangle intérieur qui n'est pas au milieu, et (4) la jonction des angles. Il est évident que l'enfant ne fait pas ordinairement une analyse aussi explicite que celle-ci, mais il faut de l'analyse même si elle est inconsciente, pour réussir.

Le succès sur ce test indiquerait donc une certaine capacité de fusionner les parties en un tout significatif. Un échec, au contraire, indiquerait que les éléments ont été perçus un par un et que la conscience est incapable de les percevoir comme un tout. Ce test se rapproche donc du test de la répétition de phrases car ce dernier permet aussi le groupement d'impressions selon leur sens, avec comme résultat, qu'il est possible de répéter cinq ou six fois plus de syllabes que de chiffres.

Binet obtenait plus d'échecs sur la grecque que sur le rectangle probablement parce qu'il plaçait le rectangle à gauche et la grecque à droite. A cause de l'influence des habitudes de lecture, il se peut fort bien qu'un bon nombre de sujets examinent tout d'abord le dessin de gauche, et à cause du peu de temps accordé (dix secondes) ne consacrent pas suffisamment de temps à celui de droite. Pour obvier à cette difficulté, Terman a placé le dessin de plus grande difficulté à gauche avec, comme résultat, un nombre plus égal d'échecs entre les deux dessins.

Les directives, les modifications et le cotation.-- Terman et Kuhlmann proposent ce test de la même façon que Binet le fait dans son échelle métrique: les directives et la façon de coter sont les mêmes. La seule modification qu'ils y apportent est le changement de place des dessins, c'est-à-dire qu'ils présentent la

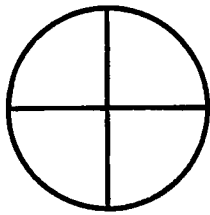
grecque à gauche et la section de prisme à droite. Nous avons vu plus haut la raison de ce changement.

Terman propose ce test à neuf ans et à onze ans dans la formule L de sa révision de 1937. Si le sujet fait un dessin ou deux demi-dessins correctement, il passe le test à neuf ans et s'il en réussit un et demi il passe celui de onze ans. Son expérience a dû lui montrer qu'obtenir un crédit et demi à dix ans était un peu trop difficile pour cet âge.

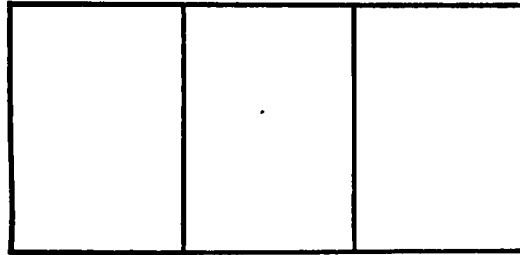
Amoss présente ce test de la même façon que Terman. Il n'y a que sa façon de coter qui diffère. Il accorde un mois de crédit pour chaque dessin reproduit correctement et un mois de crédit s'il n'y a pas plus d'une erreur dans chaque dessin. Ainsi pour obtenir le crédit total sur ce test, il faut que le sujet reproduise les deux dessins sans erreur. Il est sévère lorsqu'on le compare à Terman qui accorde les deux mois de crédit à onze ans pour une réussite et demi.

Amoss présente deux autres dessins à reproduire de mémoire à huit ans et à adulte moyen. A huit ans, il s'agit de reproduire les dessins illustrés plus bas: un cercle avec deux lignes en forme de croix à l'intérieur et un rectangle divisé en trois parties égales. Pour ce test-ci, il accorde un mois de crédit pour chaque dessin reproduit correctement, mais contrairement au test de dix ans, il n'accorde aucun crédit pour des demi-réussites.

A adulte moyen, le dessin à reproduire consiste en trois anneaux doubles entrelacés qu'on montre au sujet pendant dix secondes. Six mois de crédit sont accordés si les trois cercles sont bien réunis, quatre



Le cercle



Le rectangle

Fig. 5. Dessins à reproduire de mémoire, proposés par Amoss, à huit ans.

mois pour deux cercles bien réunis et deux mois de crédit pour un cercle bien rattaché. La difficulté de ce test vient de l'entrelacement des cercles.

Il semblerait que ce test est un bon test d'analyse et d'observation car il faut que le sujet se rende compte durant les dix secondes qu'on lui montre le dessin qu'il s'agit bien d'entrelacer les anneaux pour qu'ils soient bien reliés. En général, les sujets s'efforcent de se rappeler dans le détail que tel anneau passe par-dessus et tel autre en-dessous et ils n'ont pas le temps de l'analyser en entier.

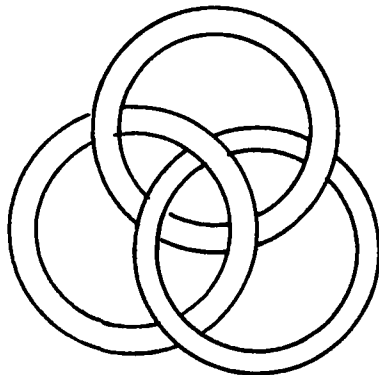


Fig. 6. Dessin à reproduire de mémoire, proposé par Amoss, à adulte moyen.

Ces deux tests sont l'invention d'Amoss. Sont-ils bien placés là où il les propose? Pour répondre, il faudrait faire une longue recherche qui est en dehors des cadres de ce travail.

Dans ce deuxième chapitre nous avons pris chaque série d'items, et après en avoir étudié les sources nous en avons fait une étude comparative et critique en relation avec ceux d'auteurs bien connus comme Terman, Kuhlmann et Wechsler. A l'exception du petit nombre, les tests ont été tirés presque textuellement de la révision Stanford-Binet de 1916. Amoss change quelques mots par-ci, par-là et il modifie la façon de coter de Terman en accordant des crédits partiels pour des tests réussis en partie seulement. Il prétend obtenir ainsi des résultats plus finement gradués, mais par le fait même il change complètement les résultats qui ne sont plus directement comparables à ceux du Stanford-Binet de Terman. Amoss ne semble pas s'en rendre compte puisqu'il emploie pour le C.I.E. la même échelle de valeur qu'emploie Terman dans le Stanford-Binet. Tant qu'une standardisation sérieuse ne sera pas faite du C.I.E., les résultats obtenus n'auront aucune valeur puisque c'est le degré de constance et de justesse d'un test qui lui donne toute sa signification dans l'interprétation des résultats obtenus.

CHAPITRE III .

ETUDE CRITIQUE ET COMPARATIVE DES ITEMS PARTICULIERS

Dans le chapitre précédent nous avons fait une étude critique et comparative des items qui se présentent en séries dans le C.I.E. Il reste encore à traiter les items particuliers qui n'apparaissent qu'une fois dans le test. Ces items se chiffrent par vingt-cinq: 1) nommer des objets familiers, 2) les formes géométriques, 3) la compréhension, 4) nommer des couleurs, 5) le jeu de patience, 6) trois commissions, 7) distinguer la droite de la gauche, 8) reconnaître des pièces de monnaie, 9) les lacunes dans les images, 10) le nombre de doigts, 11) compter à reculons de 20 à 1, 12) le coût de timbres, 13) écrire une dictée, 14) faire des phrases, 15) trouver des rimes, 16) les poids de Gesell, 17) les phrases absurdes, 18) nommer vingt mots en une minute, 19) le plan de recherche, 20) les phrases en désordre, 21) le problème d'induction, 22) renverser les aiguilles de l'horloge, 23) les problèmes d'ingéniosité, 24) la compréhension de la lecture, 25) le dessin au miroir.

Ainsi que nous l'avons fait pour les séries d'items notre étude portera sur les mêmes points de vue: 1) les sources, 2) les modifications apportées, 3) les directives, 4) le cotage proposé, et 5) la valeur des items. Tous les items ne comporteront pas une étude aussi détaillée des cinq points parce qu'un bon nombre d'items sont pris textuellement du Stanford-Binet. Connaissant la valeur de cette révision il n'y a pas lieu de douter de la validité des items lorsque Amoss n'y change rien. Nous nous bornerons donc à traiter ces derniers assez sommairement.

1. Nommer des objets familiers.

Le but de ce test est de savoir si l'enfant peut faire l'association entre des objets familiers et leur nom. Les procédés mentaux qui entrent en jeu ici sont bien élémentaires mais cependant bien fondamentaux. L'enfant normal de deux à quatre ans éprouve un grand plaisir à ce genre d'activités et un retard marqué à faire de telles associations serait une indication importante du manque de spontanéité nécessaire au développement d'une intelligence normale ou supérieure. Il ne serait pas question ici de douter de la validité de ce test sous prétexte qu'on a enseigné à l'enfant le nom des objets dont on se sert. Ce sont tous des objets très familiers. Presque tous les enfants de trois ans, quelque pauvre que soit leur milieu, se sont familiarisés avec au moins trois des cinq objets et ils ont appris leurs noms en questionnant ou en les entendant prononcer.

Ce test est tout d'abord proposé par Binet dans son échelle métrique. Il s'agit de nommer une clef, un canif et un sou. Binet place ce test à quatre ans parce qu'il faut que le sujet nomme les trois objets pour réussir cet item. Terman propose ce même test à trois ans dans sa révision de 1916, mais le sujet n'a qu'à nommer correctement trois objets sur cinq. Les objets à nommer sont une clef, un sou, un canif, une montre et un crayon. Si, comme Kuhlmann, on exigeait cinq bonnes réponses sur six ou encore trois sur trois comme le font Binet et Goddard, le test appartiendrait probablement au niveau de quatre ans. Dans le New Revised Stanford-Binet Terman propose ce même test à deux ans et six mois dans les

formules L et M de son test. Dans la formule L les objets à nommer sont une chaise, une automobile, une boîte, une clef et une fourchette. Ici, il exige quatre réussites sur cinq essais. Dans la formule M, ce test apparaît à deux ans, à deux ans et six mois et à trois ans. Les objets à nommer sont un soulier, une montre, un téléphone, un drapeau, un canif et un poêle. A deux ans Terman présente ce test comme test supplémentaire et il exige trois réussites sur six; à deux ans et six mois, quatre réussites; et à trois ans cinq réussites. Selon lui, une réussite de plus augmenterait la difficulté de l'item d'une façon assez considérable puisqu'un mot de plus suffit à placer le test à un niveau de six mois plus élevé.

Amoss prend ce test textuellement du Stanford-Binet. Des cinq objets à nommer il n'en change qu'un: au mot 'clef' il substitue le mot 'livre'. La façon de coter est la même; il exige trois bonnes réponses sur cinq. Le mot 'livre' est-il plus facile ou plus difficile que le mot 'clef'? Nous n'en savons rien. Cependant Terman prévient de ne pas changer la liste d'objets énumérés dans son test parce que, dit-il, une sélection improvisée contiendrait sûrement des objets plus ou moins familiers que ceux qui sont présentés dans la liste standardisée¹. Pourquoi Amoss a-t-il fait ce changement? Serait-ce parce qu'il voulait y ajouter une petite note personnelle? Puisqu'il se permet des modifications, il aurait bien fait de se donner la peine de chercher jusqu'à quel point cette modification change la valeur des résultats. Cependant le bon Dr Amoss ne semble pas s'inquiéter pour si peu.

1 Lewis M. Terman, The Measurement of Intelligence, Boston, Houghton-Mifflin, 1916, p. 144.

2. Les formes géométriques.

Ce test fut imaginé et standardisé par Kuhlmann pour sa révision de 1912 à l'âge de quatre ans. Il fut repris par Terman et standardisé d'une façon plus parfaite pour le Stanford-Binet. Ce dernier le gardera aussi à quatre ans dans le New Revised Stanford-Binet. Ce test comprend un carton sur lequel se trouvent dix formes géométriques et un autre carton avec les mêmes formes qu'on découpe pour les placer chacune sur la position X du premier carton. Il faut que le sujet indique du doigt sur le carton, les formes identiques à celles qui sont placées tour à tour sur la position X.

On peut considérer ce test comme une variante du form-board. Il révèle la capacité du sujet à comparer des perceptions visuelles de formes. La perception juste d'une forme même assez simple exige l'intégration d'un certain nombre d'éléments sensoriels en un tout. Les formes utilisées dans ce test ont du sens. Bien qu'un enfant de quatre ans n'ait jamais entendu parler de triangles, de carrés, etc., le sens qu'il en tire est probablement un composé de facteurs tels que l'appréciation de symétrie et de discernement de quantité.

Un élément important de succès dans ce test est sans nul doute la capacité de faire la comparaison entre la forme montrée et celles qui apparaissent sur la carte. Certains enfants se contentent d'indiquer la première forme qui leur tombe sous les yeux. Loin d'être une excuse pour leur échec, un tel manque d'attention et de sens critique serait une indication d'un niveau mental au-dessous de quatre ans. En effet, nous

avons constaté que des sujets d'intelligence inférieure n'arrivaient pas à faire ce discernement et montraient n'importe quelle forme sans même faire un effort pour les distinguer les unes des autres.

En plus de compter le nombre d'erreurs, il est très intéressant de noter avec quelles formes ces erreurs sont faites. Appareiller un cercle avec un ellipse ou un octagone, par exemple, est une erreur moins grave que d'appareiller un carré avec un triangle.

Kuhlmann, Terman et Amoss s'accordent sur la façon de présenter ce test. Ils s'en tiennent aux directives de Kuhlmann qui a imaginé ce test. Cependant Amoss diffère d'opinion avec les deux autres auteurs sur la façon de coter. Kuhlmann et Terman exigent sept réussites pour accorder les deux mois de crédit. Amoss accorde aussi deux mois de crédit pour sept réussites mais de plus il accorde un mois de crédit pour cinq réussites. Un sujet pourra obtenir ainsi un mois de crédit de plus sur le C.I.E. que sur le Stanford-Binet. Amoss se montre donc plus indulgent que ses prédécesseurs.

3. La compréhension.

Ce test facile à donner consiste tout simplement à attirer l'attention du sujet et à lui poser les trois questions suivantes: 1) Que dois-tu faire lorsque tu t'endors? 2) Que dois-tu faire lorsque tu as froid? 3) Que dois-tu faire lorsque tu as faim? On peut répéter les questions si le sujet ne répond pas.

Le but de ce test est de s'assurer si l'enfant comprend les situations qu'on lui suggère et donner une réponse assez juste. Il faut

tout d'abord qu'il comprenne le langage et qu'ensuite il dise comment faire face à la situation qu'on lui propose.

On peut se demander si l'enfant ne pourrait pas manquer de répondre aux questions correctement et cependant avoir l'intelligence nécessaire pour faire la chose appropriée quand la situation se présente. Il se peut fort bien qu'il en soit ainsi mais comme il est impossible au moment du test de lui faire avoir froid ou faim ou de lui donner sommeil pour observer sa façon d'agir, il faut donc se contenter de lui suggérer une situation qu'il doit s'imaginer. Ce travail requiert plus d'intelligence que de faire la chose lorsque l'occasion se présente.

Les deux premières questions furent proposées par Binet dans l'échelle métrique de 1905. Terman ajouta la troisième dans le Stanford-Binet. Ce test ne fut standardisé que dans la recherche Stanford de 1913-1914. Selon Terman c'est un test excellent pour quatre ans. Cependant il le présentera à trois ans et six mois dans le New Revised Stanford-Binet mais il ne proposera que la première et la dernière question et il n'exigera qu'une bonne réponse pour passer le test à cet âge.

Amoss présente et cote ce test exactement comme le fait Terman dans sa révision de 1916. Contrairement à ce qu'il fait d'ordinaire il n'accorde pas de crédit partiel pour ce test-ci. Puisqu'Amoss a jugé bon d'accorder des crédits partiels, il semblerait qu'il aurait mieux fait de suivre ce procédé pour tous les tests du C.I.E. Quelle règle, s'il en suit une, détermine si oui ou non il accordera ces crédits partiels? Nous n'en savons rien.

4. Nommer quatre couleurs.

Ce test consiste à nommer les quatre couleurs fondamentales: le rouge, le bleu, le vert et le jaune. Comme le dit Binet, il a éliminé celles dont le nom est moins familier aux enfants, le violet, par exemple, et l'orange, car son épreuve ne porte pas sur la perception et la distinction des couleurs, mais sur leur nomination². En effet, nommer les couleurs n'est pas un test de discernement des couleurs, car cette capacité se développe quelques années au-dessous du niveau où ce test-ci est placé. Tous les enfants normaux de cinq ans qui ne souffrent pas de daltonisme peuvent discerner les quatre couleurs primaires aussi bien que n'importe quel adulte.

La capacité de faire des associations simples entre une impression sensitive et un objet est acquise avant celle des associations de couleurs. Plusieurs objets peuvent être nommés par l'enfant de trois ans. Il peut paraître étrange à première vue que le nom des couleurs soit appris beaucoup plus tard mais il faut se rappeler que l'enfant n'a pas d'occasions aussi nombreuses d'observer les couleurs et d'entendre leur nom. De plus, la désignation des couleurs par leur nom n'a pas d'ordinaire beaucoup de valeur pratique pour le jeune enfant. Lorsqu'il les apprend, c'est surtout à cause de son intérêt spontané pour le monde des sens. Ce manque d'intérêt à cinq ans est toujours un mauvais signe et il serait une indication de retard dans le développement de l'enfant. L'imbécile n'arrive

² Alfred Binet et Thomas Simon, La mesure du développement de l'intelligence chez les jeunes enfants, Paris, Société Alfred Binet, 1938, p. 76.

jamais à associer les couleurs avec leur nom. Les résultats des filles seraient quelque peu supérieurs à ceux des garçons, sans doute à cause de leur intérêt plus prononcé pour les couleurs³.

Binet plaça tout d'abord ce test à huit ans dans son échelle de 1905, mais il le descendit à sept ans dans celle de 1911. Goddard le place aussi à sept ans tandis que Kuhlmann, Terman et Amoss le proposent à cinq ans. Terman nous dit que toutes les statistiques justifient sa place à cinq ans. Les données de Bobertag sont l'exception auxquelles s'opposent celles de Rowe, Winch, Dumville, Dougherty, Brigham et les trois recherches Stanford⁴.

Dans les directives que Terman propose, il spécifie bien de poser la question suivante: "Quel est le nom de cette couleur?" et non pas "Quelle couleur est ceci?" parce que, dit-il, cette formule pourrait provoquer la réponse: "La première couleur" ou encore "une jolie couleur" et il ne permet pas de revenir et de poser la question de nouveau après qu'une couleur a été mal nommée. Le test est réussi si l'enfant nomme toutes les couleurs correctement.

Dans ses directives Amoss propose les questions suivantes: "Quelle couleur est ceci?" ou encore "Quel est le nom de cette couleur?" La première question est justement celle à laquelle Terman s'oppose puisqu'elle est équivoque et peut suggérer une autre réponse que celle que l'on désire. C'est donc là une première lacune. Amoss nous dit encore que si le sujet

3 Terman, op. cit., p. 165.

4 Loco cit.

ne répond pas correctement la première fois, l'examineur renverse la carte et répète le test. C'est donner une deuxième chance au sujet. Si dans un des deux essais le sujet nomme toutes les couleurs correctement on lui accorde les deux mois de crédit, si dans un des essais il parvient à nommer trois couleurs correctement on lui accorde un mois de crédit.

Il est évident que selon les directives d'Amoss la valeur de ce test est changée complètement et elle n'est plus comparable à celle de Terman. Il y a donc lieu de douter fortement que ce test ne soit pas bien placé dans l'échelle d'Amoss car nous avons vu dans un test précédent que réussir tous les items d'un test augmente de beaucoup la difficulté, et de la façon qu'il est proposé par Terman ce test est beaucoup plus difficile que celui d'Amoss. Non content de cela, Amoss en diminue encore la difficulté en accordant un mois de crédit pour trois réussites dans un ou l'autre essai. Ainsi coté le test pourrait être placé plus bas dans l'échelle.

Selon Terman⁵ ce test semblerait plus influencé par l'entourage que n'importe quel autre test de l'échelle et si le niveau social de l'enfant est bas un échec ne serait significatif qu'après l'âge de six ans. C'est sans doute la raison pour laquelle Terman l'omet dans sa révision de 1937.

5 Loco cit.

5. Le jeu de patience.

Le jeu de patience se compose de deux cartes rectangulaires dont une est coupée en diagonale formant deux triangles. Le rectangle est placé sur la table devant le sujet et il faut qu'avec les deux triangles il forme un rectangle semblable à celui qui est placé sur la table.

Comme le décrit Binet, l'opération comporte les éléments suivants:

1) il faut que le sujet se souvienne du but à atteindre, c'est-à-dire la figure qu'il doit former, 2) il doit essayer différentes combinaisons sous l'influence de l'idée dirigeante qui guide ses efforts même s'il n'en a pas conscience, 3) il doit porter un jugement sur ce qu'il a réalisé, le comparer avec le modèle et décider s'il a réussi⁶.

On peut donc ranger cet item dans la classe de tests dits de la combination method, puisque les éléments doivent être combinés en un tout selon une idée dirigeante. Sous ce rapport il a quelque chose de commun avec le form-board, le test d'Ebbinghaus et les phrases à mots mélangés. Binet le désigne comme un "jeu de patience" parce que le succès dépend d'une certaine bonne volonté à persister dans une ligne de conduite sous l'influence d'une idée dominante⁷.

Tous les échecs dans ce test n'auront pas la même signification. Un enfant de cinq ans même s'il est très intelligent peut quelquefois échouer, mais la plupart du temps pas sans avoir essayé plusieurs combinaisons qu'il rejette ensuite comme n'étant pas satisfaisantes. L'enfant

6 Binet et Simon, op. cit., p. 171.

7 Loco cit.

d'intelligence inférieure s'arrêtera souvent après avoir juxtaposé les morceaux d'une façon quelconque ou même absurde, et il se contentera de cet effort parce qu'il est borné et n'a aucune sens critique.

Ce test proposé tout d'abord par Binet semble trop facile pour cinq ans et un peu trop difficile pour quatre ans. C'est la façon de présenter le test qui le rend trop facile. Selon les directives de Binet, les deux triangles doivent être disposés de telle manière que les deux hypothénuses forment un angle droit l'une avec l'autre tel qu'illustré dans la figure 6.

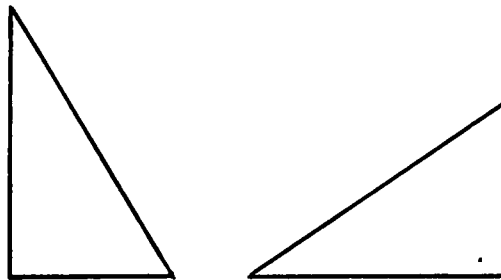


Fig. 6. (a) Position des triangles préconisée par Binet pour le jeu de patience.

Le sujet n'a qu'à glisser les deux cartons en refermant l'angle pour réunir les deux morceaux correctement.

Terman augmente la difficulté du test en disposant les deux triangles à un pouce de distance, de façon à ce que les deux hypothénuses soient tournées à l'extérieur tel qu'illustré dans la position (a) de la figure 7. Le sujet doit donc transposer les deux morceaux (position b) et ensuite tourner le rectangle pour que le côté le plus long soit parallèle à la table (position c). Le sujet doit réussir deux essais sur trois.

Kuhlmann propose et cote ce test de la même façon que le fait Terman

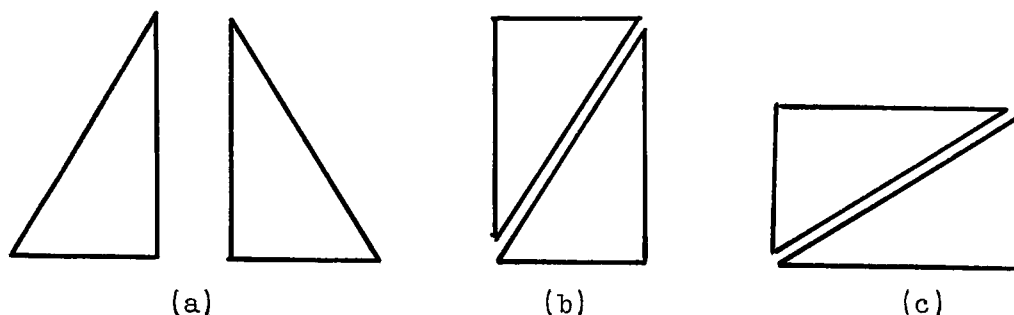


Fig. 7. Position des triangles préconisée par Terman pour le jeu de patience.

Voyons maintenant ce que fait Amoss. Il adopte une autre disposition. Les deux cartons sont placés devant le sujet à un pouce de distance de façon à ce que les deux côtés les plus courts soient parallèles à la ligne de vision du sujet tels qu'illustrés dans la position (a) de la figure 8. Le sujet doit réussir deux essais sur trois.

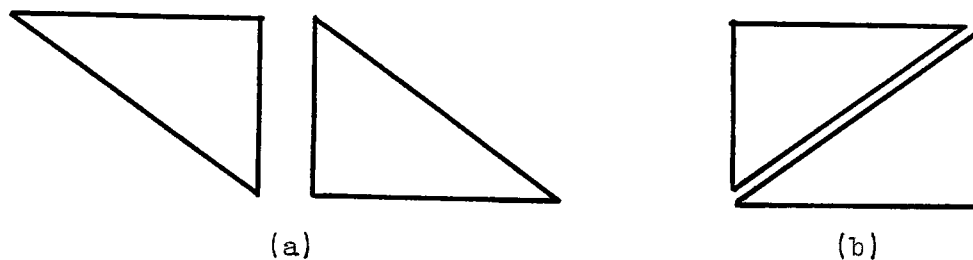


Fig. 8. Position des triangles préconisée par Amoss pour le jeu de patience.

Cette disposition proposée par Amoss rend le test beaucoup plus facile que celle de Terman parce que le sujet n'a qu'à faire une opération c'est-à-dire renverser la position des deux triangles. Le fait que les côtés plus longs des triangles soient placés horizontalement comme le sont d'ailleurs ceux du rectangle facilite la vue d'ensemble. Il n'y a qu'à considérer un moment la figure 8 pour constater qu'il est beaucoup plus facile de résoudre le problème tel que proposé par Amoss.

Un léger changement dans les directives d'un test peut paraître banal et sans conséquence pour quelqu'un qui n'est pas initié à la méthode des tests; cependant un pareil changement suffit pour rendre le test non valide. Si on fait confiance aux statistiques de Terman pour le jeu de patience, la présentation d'Amoss le rendrait trop facile pour cinq ans.

6. Exécuter trois commissions.

Binet nous dit⁷ que c'est la connaissance des faits suivants qui lui ont inspiré l'épreuve des trois commissions. On demande de bonne heure aux petits enfants de faire des commissions chez le marchand. Les médecins qui ont l'habitude de la clinique des arriérés savent que ces enfants se reconnaissent non pas à ce qu'ils sont incapables de faire une commission mais plutôt incapables d'en faire plusieurs à la fois. Alors Binet propose cette épreuve qui consiste à faire faire trois commissions à l'enfant: mettre une clef sur une chaise qu'on lui désigne, fermer ou ouvrir une porte, et apporter à l'examineur une boîte qui se trouve sur une chaise. On dit au sujet ce qu'il doit faire après quoi on lui demande s'il a bien compris et on répète une autre fois les directives. On doit donner les commissions toujours dans le même ordre et on ne doit plus répéter les directives une fois qu'elles ont été données au complet. Il faut que les trois commissions soient exécutées dans l'ordre donné pour réussir le test.

7 Binet et Simon, op. cit., p. 75.

Le succès dépend tout d'abord de la capacité de comprendre les commandements et ensuite de la capacité de les retenir. C'est donc en partie un test de mémoire bien que ce soit d'une sorte différente de celle de la répétition de chiffres ou de phrases. Terman croit que c'est un test excellent puisqu'il jette de la lumière sur un aspect de l'intelligence qui est nécessaire dans n'importe quelle occupation et dans la vie de tous les jours⁸.

Binet plaça tout d'abord ce test à six ans mais il le monta à sept ans dans sa révision de 1911. Les trois recherches Stanford et les statistiques de tous les chercheurs témoignent sans aucun doute que ce test est bien placé à cinq ans⁹. Cependant Terman placera ce test à quatre ans et six mois dans la formule L du New Revised Stanford-Binet.

Amoss propose ce test à cinq ans et il le cote de la même façon que Terman. Il ne fait que changer les objets: mettre un crayon sur la table, ouvrir ou fermer une porte, et apporter à l'examineur un livre qui se trouve sur une chaise. Il semble peu probable que ces légères modifications changent la valeur du test puisque les objets qu'il propose sont également familiers à l'enfant, l'important étant de se rappeler l'ordre des commissions. Il n'y a donc pas lieu de douter que la valeur du test ait été modifiée par ces petits changements.

8 Terman, op. cit., p. 75.

9 Loco cit.

7. La droite et la gauche.

Pour savoir si l'enfant fait la distinction entre la droite et la gauche on lui pose les questions suivantes: "Montre-moi ta main droite." "Montre-moi ton oeil gauche," et enfin "Montre-moi ton oreille droite." Si le sujet fait une erreur, on recommence avec la main gauche, l'oeil droit et l'oreille gauche. Le test est réussi si l'enfant donne trois bonnes réponses sur trois ou cinq sur six. Si l'enfant se corrige, on tient toujours compte de la deuxième réponse que cette dernière soit bien ou mal.

Il est intéressant de suivre chez l'enfant l'acquisition de distinctions dans le langage se rapportant à l'orientation dans l'espace. Autres distinctions de ce genre sont celles entre en haut et en bas, au-dessus et au-dessous, près et loin, en avant et en arrière, etc. Comme l'a fait remarquer Bobertag¹⁰ l'enfant maîtrise des distinctions telles qu'en haut et en bas, au-dessus et au-dessous, en avant et en arrière avant la droite et la gauche. Terman avance quatre théories pour expliquer ce retard dans la distinction de la droite et de la gauche comparée à celle d'en haut et d'en bas. Il s'expliquerait 1) par la fréquence avec laquelle les enfants ont l'occasion de faire les distinctions respectives, 2) par la supposition que les sensations de kinesthésie sont plus marquées dans la distinction d'en haut et d'en bas que de gauche et de droite, 3) par le fait qu'un objet vu tantôt d'un côté et tantôt de

10 Terman, op. cit., p. 176.

l'autre montrent deux aspects qui se ressemblent de près et s'identifient presque dans l'idée de l'enfant tandis que les aspects correspondant à en haut et en bas demeurent plus distincts dans l'esprit de l'enfant, 4) par le langage puisque l'enfant entend plus souvent en haut et en bas que gauche et droite il en apprend le sens plus vite¹¹. Cette dernière explication semble très plausible mais il est fort probable qu'il y a d'autres facteurs qui entrent en jeu.

Tous les jours on peut constater que même des adultes intelligents ont de la difficulté à distinguer la gauche de la droite. Si on pose la question: "Quelle est votre main droite?" un certain nombre d'adultes hésiteront et prendront quelques secondes avant de répondre. Certains chercheront une marque distinctive qu'ils savent être sur la main droite, d'autres feront le mouvement d'écrire parce qu'ils savent qu'ils écrivent de la main droite. Ces mêmes personnes cependant n'ont aucune difficulté à distinguer en haut et en bas.

Considérant cette difficulté qu'on trouve chez des adultes tout à fait normaux, on peut se demander s'il est juste de s'en servir comme test d'intelligence. Terman répond par l'affirmative parce que, dit-il, des adultes normaux, malgré la confusion momentanée, peuvent invariablement faire la distinction, si ce n'est par association directe, au moins par association indirecte. Les parfaits imbéciles, quel que soit leur âge, ne parviennent jamais à passer ce test¹².

11 Terman, op. cit., p. 176.

12 Terman, op. cit., p. 178.

Binet plaça ce test à six ans dans son échelle de 1908, mais il le monta à sept ans dans celle de 1911. Les statistiques des recherches Stanford aussi bien que les données de tous les auteurs à l'exception de celles de Bobertag justifient sa retention à six ans¹³.

Amoss copie textuellement Terman de sorte qu'il n'y a pas lieu de douter de la validité de cet item.

8. Reconnaître des pièces de monnaie.

Certains critiques de l'échelle métrique considèrent ce test comme ayant peu de valeur parce qu'ils prétendent que la capacité d'identifier des pièces de monnaie dépend entièrement de l'instruction acquise et du milieu. Les statistiques de Terman¹⁴ montrent qu'il est très peu influencé par des différences d'entourage, bien que les enfants de milieux pauvres réussissent un peu mieux que les enfants de milieux riches et cultivés. Il semble que tous les enfants de six ans ont eu l'occasion d'apprendre le nom des pièces de cinq, dix et vingt-cinq sous et que s'ils ne réussissent pas à les identifier c'est plutôt par un manque d'intérêt qui présupposerait un manque de vivacité d'esprit.

Ce test apparaît tout d'abord dans l'échelle métrique de 1908 à sept ans mais il est omis de la révision de 1911. Kuhlmann le retient à sept ans mais il exige que le sujet nomme les quatre pièces correctement. Terman le propose à six ans mais seulement trois des quatre pièces doivent être identifiées, le rendant ainsi assez facile pour cet âge.

13 Loco cit.

14 Terman, op. cit., p. 184.

Le seul changement qu'Amoss fait à ce test, c'est d'exiger une réponse dans l'espace de vingt secondes. Il semble bien que cette limite de temps influence très peu les résultats. Si le sujet est incapable de donner la réponse en vingt secondes, c'est qu'il ne la connaît pas et il est peu probable qu'il répondrait correctement si un temps plus long lui était accordé. Il n'y a donc pas lieu de douter que cette modification change beaucoup la valeur du test. Toutefois, Terman l'éliminera dans sa révision de 1937 sans doute pour une raison de validité.

9. Les lacunes dans les images.

Ce test est une des nombreuses variantes du completion test qui ont toutes en commun le trait suivant: il faut trouver des lacunes dans les parties d'un tout. Le tout à compléter peut être un mot, une phrase, une histoire, une image, un objet, etc. Quelquefois, toutes les parties du tout sont données et il ne faut que les mettre en ordre comme, par exemple, les phrases en désordre.

Bien qu'il y ait une certaine ressemblance dans tous ces tests, ce ne sont pas toujours les mêmes procédés mentaux qui entrent en jeu. Le facteur principal peut être la cohérence du langage, ou bien la perception visuelle des formes ou encore l'association d'idées abstraites.

Dans ce test-ci, il faut tout d'abord que les parties soient perçues comme un tout et ensuite que l'idée de l'objet puisse être reproduite assez facilement et assez clairement pour qu'elle serve de modèle pour la critique de l'image en question. Plus l'enfant est jeune, moins cette faculté de perception est développée. Ce n'est pas avant six ans que l'enfant peut

analyser les images et trouver ce qui manque. Les perceptions des jeunes enfants sont très peu critiques et très peu analytiques, mais elles le deviennent avec l'âge et l'expérience. Le jeune enfant dessine un homme en lui faisant une tête, des bras et des jambes et il en sera complètement satisfait. L'accroissement de son développement mental se manifestera par l'addition des détails qui manquent tels que le tronc, les cheveux, le nez, la bouche, etc. On peut aussi constater qu'un enfant peut identifier et nommer une voiture dans une image mais ce ne sera que bien plus tard qu'il remarquera l'absence d'une roue malgré qu'il puisse identifier une roue aussi bien qu'une voiture.

En standardisant une série d'images absurdes Terman¹⁵ constata que des enfants normaux de trois ans ne voyaient rien d'anormal à une image qui montrait un chat à deux pattes ou une poule à quatre pattes. Cependant ces mêmes enfants ne prendraient jamais une poule pour un chat. La difficulté vient de leur incapacité de se faire une image suffisamment claire pour servir de comparaison à celle qu'ils ont sous les yeux. Les imbéciles ont la même difficulté que les jeunes enfants à reconnaître des lacunes dans des images d'objets familiers.

Binet plaça tout d'abord ce test à sept ans pour le monter à huit ans dans sa révision de 1911. Il a été retenu à sept ans dans presque toutes les révisions de l'échelle métrique bien que toutes les statistiques à l'exception de celles de Bobertag justifient sa place à six ans¹⁶.

15 Terman, op. cit., p. 180.

16 Loco cit.

Dans sa révision de 1916 Terman présente trois figures auxquelles manquent un oeil, le nez, la bouche, et un mannequin auquel il manque les bras. Il faut que le sujet trouve trois des quatre lacunes pour réussir le test. Dans la formule L de la révision de 1937 il en augmente la difficulté en proposant cinq objets: une voiture, une paire de bottines, une théière, un lapin et un gant auxquels il manque respectivement une roue, un lacet, un anse, une oreille et un doigt. Il faut que le sujet réussisse quatre des cinq items pour passer le test.

Amoss présente quatre figures auxquelles il manque un oeil, une oreille, le nez et la bouche. Les figures sont beaucoup plus grosses que celles de Terman de sorte que les lacunes sont beaucoup plus évidentes. L'enfant n'a qu'à susciter l'image d'une figure dans son esprit et il a toujours celle de l'expérimentateur pour servir de modèle. C'est peut-être la raison pour laquelle Terman change les figures pour des objets qui ne sont pas présents à l'enfant au moment du test.

Il est évident que la difficulté de ce test sera diminuée ou augmentée selon le choix des images et des lacunes qu'il faut y trouver. Comme pour chacun des tests d'ailleurs, il n'y a que la standardisation pour nous dire si cet item est bien placé à l'âge où il est proposé. Il semble bien que Terman l'a trouvé trop facile puisqu'il y apporte d'importantes modifications qui en augmentent la difficulté dans sa révision de 1937.

10. Le nombre de doigts.

Binet présente ce test à sept ans dans sa révision de 1908, mais il l'omet dans celle de 1911. Kuhlmann et Terman le gardent à sept ans où il appartient d'après leurs statistiques.

Comme pour compter des sous, ce test-ci jette aussi de la lumière sur l'intérêt de l'enfant pour les nombres. Cependant, les procédés mathématiques qui entrent en jeu ici sont plus compliqués que ceux qui sont nécessaires pour compter des sous. Si l'enfant peut donner le nombre de doigts, c'est parce qu'il les a déjà comptés et qu'il se souvient du résultat. Il ne s'en souviendrait pas à moins d'avoir un certain intérêt pour les nombres. Des imbéciles d'âge adulte se souviennent rarement du nombre de leurs doigts, même après leur avoir répété maintes et maintes fois. Ils sont incapables de former des concepts précis autres que les relations les plus simples des nombres parce qu'ils y trouvent peu d'intérêt ou de signification.

Terman¹⁷ a constaté que ce test était très peu influencé par la scolarité ou le milieu social. Ses données prouvent qu'il remplît toutes les conditions d'un bon test. Il y a une augmentation rapide dans le pourcentage des sujets qui réussissent ce test de cinq à huit ans. Tous les auteurs s'accordent sur la valeur véritable de ce test.

La façon de donner le test et de le coter est la même pour tous les auteurs. Il faut que la réponse soit donnée sans hésitation et ils n'accordent que trois secondes. Seul, Amoss accorde dix secondes bien

17 Terman, op. cit., p. 110.

qu'il prévienne que le sujet doit être averti s'il commence à compter ses doigts ouvertement ou furtivement. Il est évident qu'un sujet qui prend de cinq à dix secondes pour répondre a pris le temps de compter et il peut fort bien le faire sans que l'examineur puisse s'en rendre compte autrement que par sa lenteur à répondre. Cependant l'examineur devra accorder les points si la réponse est donnée en dix secondes.

Pourquoi cette mise en garde pour essayer d'empêcher le sujet de compter si d'un autre côté l'examineur doit lui donner amplement le temps de le faire? C'est un non-sens. Amos rend donc ce test d'une validité très douteuse en accordant dix secondes au sujet pour donner sa réponse.

11. Compter à reculons de vingt à un.

Compter à reculons de vingt à un est un autre test tiré de l'échelle Binet-Simon de 1908.

On peut poser l'objection que compter à reculons est un test de scolarité, et c'est vrai jusqu'à un certain point. Il est vraisemblable que des exercices spéciaux rendraient un enfant capable de passer ce test un peu plus tôt qu'à l'ordinaire bien qu'il soit douteux que des enfants en bas de sept ans aient eu suffisamment d'entraînement pour influencer leurs résultats. D'autre part, lorsqu'un enfant a atteint le développement mental de huit ans et au plus neuf ans, il est capable de compter de vingt à un même s'il n'a jamais essayé de le faire.

Ce test présuppose la capacité de compter de un à vingt. Ceci ne garantit pas toutefois le succès pour compter à reculons. Il faut plus

que la mémoire des noms des chiffres jusqu'à vingt; il faut que l'enfant soit capable d'une attention soutenue jusqu'à ce que sa tâche soit terminée. Il doit toujours avoir présent à l'esprit le fait qu'il lui faut aller à reculons. Si l'enfant perd de vue cette fin, il peut fort bien être subjugué par son habitude de compter à la suite et, rendu à treize ou quatorze, continuer à la suite. On peut donc considérer ce test comme un test d'attention ou de contrôle prolongé de la pensée. Cette capacité d'exercer une vigilance continue durant une période de vingt à trente secondes se rencontre rarement au-dessous d'un niveau mental de sept ou huit ans.

Les statistiques sur ce test s'accordent très bien. Il est évidemment trop facile à neuf ans et trop difficile à sept ans. Tous les auteurs le placent à huit ans. Le manque d'uniformité vient dans la limite de temps fixée pour la répétition. Binet n'accorde que vingt secondes; Kuhlmann en accorde trente pour une répétition sans erreur et vingt pour une répétition avec une erreur. Terman nous dit¹⁸ qu'il est d'avis avec Bobertag qu'il ne faut pas être trop sévère sur ce point. Bien que la plupart des enfants puissent compter à reculons en vingt secondes, il y a des sujets intelligents mais réfléchis qui prennent de trente à quarante secondes. Si le sujet compte avec assurance ces deux auteurs ne voient pas d'inconvénient à accorder quarante secondes. Terman prévient toutefois qu'il ne faut pas être plus indulgent car un

18 Terman, op. cit., p. 214.

temps plus long permettrait au sujet de compter à la suite chaque fois qu'il veut obtenir le chiffre suivant.

Amoss diffère d'opinion sur ce point. Dans une note au bas de cet item il dit:

This being a test of mental power and not of achievement, no time limit has been set. A candidate who has no experience in counting backward will require considerably more time than one of no higher intelligence who is somewhat practiced in the operation ¹⁹.

Il a raison mais seulement jusqu'à un certain point. Il faut une limite car autrement rien n'empêche le sujet de compter à la suite chaque fois pour trouver le chiffre suivant et s'il parvient à dire tous les chiffres à reculons de cette façon, on ne peut tout de même pas prétendre qu'il peut compter à reculons. Amoss seul rend donc ce test de validité douteuse en n'imposant pas de limite de temps.

12. Le coût de timbres.

Ce test est présenté par Binet dans son échelle de 1908 à huit ans, mais sous une forme différente de celle présentée par ses successeurs. Le but que Binet se propose est de savoir si un enfant de six ans peut trouver la somme de trois sous simples et de trois sous doubles.

Terman modifie ce test en faisant usage de timbres parce que la monnaie américaine ne se prête pas au test de la façon qu'il est présenté par Binet. Il présente un carton sur lequel sont collés trois timbres

¹⁹ Harry Amoss and Charles Stogdill, Canadian Intelligence Examination, Toronto, The Ryerson Press, 1940, p. 47.

d'un sou et trois timbres de deux sous arrangés comme suit: 111222. Ce test est placé à neuf ans parce que l'auteur exige que l'enfant connaisse la valeur des timbres et qu'il en donne le total. Si le sujet connaît la valeur des timbres mais qu'il donne une mauvaise réponse pour le total, il permet un second essai. Terman omettra ce test dans sa révision de 1937.

Kuhlmann présente ce test à huit ans et à la façon de Binet. Si le sujet ne connaît pas la valeur des timbres on le lui dit et il ne s'agit que d'en trouver le total. Présenté ainsi ce test est bien placé à huit ans.

Amoss imite Binet et Kuhlmann en plaçant ce test à huit ans. Il utilise trois cartes sur lesquelles sont collés quatre timbres de deux sous sur la première, trois timbres de trois sous sur la seconde et trois timbres d'un, de deux et de trois sous sur la dernière. L'examineur fait connaître au sujet la valeur des timbres et ce dernier n'a qu'à donner la valeur totale des timbres sur chaque carte.

Amoss accorde les deux mois de crédit si le sujet donne la bonne réponse aux trois questions dans l'espace de trente secondes pour chacune, et un mois de crédit s'il donne deux bonnes réponses. La limite de temps et la façon de coter, c'est-à-dire exiger trois réponses sur trois, augmentent la difficulté du test mais par contre, le fait d'avoir trois timbres de même valeur sur les deux premières cartes en diminue la difficulté. Présenté ainsi cet item est-il bien placé à huit ans? Il semblerait qu'il l'est, mais il n'y a que la standardisation qui nous le dira. Au fond, qu'on suive un procédé ou l'autre, peu importe, pourvu que les

statistiques de la standardisation prouvent que tel qu'il est proposé le test est bien placé au niveau d'âge où il est présenté. C'est le point qu'on ignore toujours dans le C.I.E.

13. Ecrire une dictée.

Binet ne propose pas ce test dans son échelle de 1905, mais il le présente à huit ans dans celle de 1908 pour le supprimer dans celle de 1911. Il a été omis de toutes les révisions excepté celle de Terman en 1916. Bobertag considérait qu'il ne méritait même pas d'être mis à l'épreuve. La critique générale c'est qu'il est un test de scolarité plutôt qu'un test d'intelligence. Il est évident que ce test dépend de la scolarité du sujet. Si un enfant de huit ans n'a jamais appris à écrire, il ne parviendra jamais à passer ce test quel que soit son degré d'intelligence. Terman nous dit qu'un enfant de huit ans qui a un âge mental de six ans peut réussir ce test mais qu'il est peu probable qu'un enfant qui possède une mentalité au-dessous de six ans parvienne à passer ce test après deux ans de scolarité²⁰. Le test perd toute sa valeur diagnostique pour un niveau mental au-dessus de huit ans puisqu'on peut montrer à écrire aux arriérés qui ont un niveau intellectuel de huit ou neuf ans. L'expérience démontrera à Terman le peu de valeur de ce test puisqu'il le supprimera dans sa révision de 1937.

Amoss, comme Terman, propose ce test comme test supplémentaire à huit ans. La phrase à écrire est la même: See the little boy.

20 Terman, op. cit., p. 232.

Cependant Amoss se montre plus sévère que Terman pour le coter. Ce dernier accorde le crédit de deux mois même s'il y a des fautes d'épellation pourvu que la phrase soit lisible et sans omissions. Amoss n'accorde les deux mois de crédit que si la phrase est écrite sans omissions et sans fautes. Il accorde un mois de crédit s'il n'y a pas plus d'une faute d'épellation. Là où un sujet recevrait deux mois de crédit sur le Stanford-Binet, il ne recevrait qu'un mois sur le C.I.E.

Comme nous l'avons vu plus haut ce test n'a pas beaucoup de valeur comme test d'intelligence et Amoss aurait mieux fait de l'omettre.

14. Faire des phrases.

Terman nous dit²¹ que Masselon fut le premier à se servir du test de constructions de phrases contenant des mots donnés et on l'a appelé "l'expérience de Masselon". Meumann qui s'en sert dans une recherche plus élaborée²² trouve que c'est un bon test d'intelligence et un index juste de la richesse, de la précision et de la maturité des procédés d'association. Comme l'indique Meumann, il est très intéressant d'étudier les différences qualitatives entre les réponses d'enfants brillants et celles d'enfants arriérés. Elles se distinguent surtout par la logique des associations et par la précision de l'exposé. Les enfants brillants se serviront davantage des mots comme des clefs de voûte dans la cons-

21 Terman, op. cit., p. 245.

22 Lewis M. Terman, The Measurement of Intelligence, Boston, Houghton-Mifflin, 1916, p. 245 citant Meumann Ueber eine neue Methode der Intelligenzprüfung und über den Wert der Kombinationsmethoden, dans Zeitschrift für Pädagogische Psychologie und Experimentelle Pädagogik, 1912, p. 145 - 163.

truction d'une phrase qui leur sera suggérée logiquement par ces mêmes mots. La phrase de l'enfant lent, bien que correcte dans sa structure et suffisamment exempte d'absurdités, exprimera plutôt des idées qui sont plus ou moins difficiles à classer parce qu'elles ne sont pas suggérées logiquement par les mots.

Ce test est une des nombreuses formes du completion test. Comme nous le faisons remarquer dans le test des lacunes à trouver dans des images, la capacité de combiner des éléments séparés et isolés en un tout logique est un des aspects les plus essentiels de l'intelligence. L'habileté de faire ceci dans un cas déterminé dépend tout d'abord du nombre et de la qualité logique des associations qui ont été faites dans le passé avec chacun des éléments pris séparément et ensuite de la facilité avec laquelle ces idées produisent les associations particulières qui sont nécessaires pour donner de l'unité. Il est donc facile de comprendre que des sujets qui ont un bagage bien mince d'idées et dont les associations sont basées sur l'accidentel plutôt que sur la logique, trouvent ce test difficile.

Binet propose ce test à dix ans; Goddard le garde à ce niveau tandis que Terman et Kuhlmann lui assignent le niveau de neuf ans. Binet ne donne qu'un essai et se sert de mots difficiles: 'Paris, rivière, fortune'. Il augmente encore la difficulté du test en exigeant que le sujet écrive la phrase. Certains auteurs ont suivi de près le procédé de Binet et n'ont fait que substituer à 'Paris' le nom d'une ville plus connue.

Kuhlmann introduit des variantes. Dans son test l'examineur demande au sujet de faire une phrase avec les trois mots: 'argent, rivière, St-Paul'. Si le sujet ne répond pas ou ne semble pas comprendre, l'examineur doit illustrer de la façon suivante après l'espace d'une minute: "Si je te donnais les mots: 'printemps, fleurir et fleurs', (les mots s'utilisent mieux en anglais: spring, bloom, flower), tu pourrais dire: 'Les fleurs fleurissent au printemps'. Maintenant, sers-toi des mots 'argent, rivière, St-Paul' de la même façon dans une phrase." On n'accorde alors qu'une demi-minute.

Si dans l'espace de la première minute le sujet donne les mots dans trois phrases séparées ou dans trois phrases réunies par 'et', l'examineur lui dit: "Bien. Maintenant, vois si tu peux mettre ces trois mots dans une seule phrase," et on lui accorde trente secondes.

Si le sujet donne une bonne réponse sans la moindre hésitation ou encore qu'il y a raison de supposer qu'il a été préparé au test, l'examineur peut se servir des mots 'arbres, maison, travail', ou encore 'été, homme, pluie' au lieu des précédents. Le test est réussi si l'enfant donne une bonne phrase avec les trois mots pourvu qu'il n'y ait pas plus de deux propositions et qu'elle soit donnée dans l'espace d'une minute et demie.

Terman ne présente pas ce test de la même façon. Voici les directives qu'il propose: "Tu sais ce que c'est qu'une phrase. Une phrase est faite de mots qui disent quelque chose. Maintenant, je vais te donner trois mots et tu dois me faire une phrase dans laquelle tu mettras les trois mots: 'garçon, balle, rivière'. Maintenant, fais une phrase dans

laquelle se trouvent ces trois mots." On fait de même avec les deux autres groupes de mots 'travailler, argent, hommes' et 'désert, rivières, lacs'. Si l'enfant a de la difficulté à comprendre on lui explique ainsi: "On doit mettre ces mots avec d'autres mots pour que tous ensemble ils fassent une phrase." On ne parle pas de temps mais si le sujet ne donne pas la réponse dans l'espace d'une minute on considère cet item particulier comme un échec. Il faut deux réussites sur trois pour réussir le test. Terman donne ensuite une longue liste d'exemples de phrases qui sont satisfaisantes et de phrases qui ne le sont pas. Il fait remarquer que le grand obstacle à l'uniformité dans l'usage de ce test vient de la difficulté à coter, et surtout lorsqu'il s'agit de décider si la phrase est suffisamment absurde pour être classée comme un échec. Il espère que la longue liste d'exemples réduira ces difficultés au minimum.

Dans sa révision de 1937, il présente ce test à sept ans mais alors il le simplifie beaucoup. L'examineur dit au sujet: "Je vais te dire quelque chose à propos de chien et chat: 'Le chien court après le chat'. Maintenant fais-moi une phrase avec 'cheval, plus gros, chien', 'garçon, tombe, jambe', et 'enfant, fleurs, champ'." Si l'enfant hésite, l'examineur l'encourage en disant: "Je t'ai dit quelque chose à propos de chien et chat, maintenant dis-moi quelque chose à propos de . . ." Il faut deux bonnes réponses sur trois pour réussir. Il est évident que cette présentation et les mots choisis sont beaucoup plus faciles que ceux de sa révision précédente mais il est à remarquer qu'il place ce test à sept ans.

Voici maintenant la façon dont Amoss présente ce test à neuf ans. L'examineur dit au sujet: "Tu as lu des phrases dans ton livre de lecture.

Tu as vu ton maître écrire une phrase au tableau noir. Une phrase est une petite histoire comme celle-ci: 'L'homme met l'argent dans sa poche'. Je vais te donner trois mots et je veux que tu me fasses une phrase dans laquelle se trouvent les trois mots: 'garçon, rivière, balle'. Fais une phrase dans laquelle se trouvent ces trois mots. Maintenant fais une phrase avec ces trois mots-ci: 'chaise, fille, cuisine', et encore 'livre, page, non'."

A chaque fois les mots doivent être donnés lentement et distinctement mais ils ne doivent pas être répétés. Si le sujet donne trois bonnes phrases qui font du sens et qui ne contiennent pas plus de deux propositions coordonnées, on accorde les deux mois de crédit; si le sujet donne deux bonnes phrases on accorde un mois de crédit.

Amoss rend ce test beaucoup plus difficile en ne permettant pas de répéter les mots une fois qu'ils ont été donnés. Le test n'est donc plus seulement un test pour savoir si l'enfant peut faire des phrases avec des mots mais aussi un test d'attention. Il nous semble que cette interdiction n'a pas sa raison d'être dans un test de ce genre. Le test est encore rendu plus difficile par sa façon de coter. Exiger trois bonnes réponses sur trois augmente beaucoup la difficulté d'un test. Là où un sujet obtiendrait deux mois de crédit sur le Stanford-Binet il n'en obtiendrait qu'un sur le C.I.E. Il n'y a que le premier groupe de mots qui sont les mêmes que ceux de Terman; Amoss présente d'autres mots pour les deux autres groupes.

Nous avons vu que Binet, Terman, Kuhlmann et Amoss présentent ce test chacun à leur façon en le modifiant à leur gré. Qui a raison? Quelle

est la valeur de ce test présenté de façons si différentes. Au fond, ce n'est pas la présentation qui importe car chacun auteur est libre d'y apporter les modifications qu'il désire pourvu que les statistiques de sa standardisation prouvent que le test est bien placé à l'âge où l'auteur le propose. Nous savons que Terman et Kuhlmann ont fait des standardisations sérieuses. Il n'y a donc pas lieu de douter que malgré les variantes de ces deux auteurs le test soit bien placé là où il a été proposé. Mais comme pour tous les autres tests qu'il change à son gré, Amoss s'est-il donné la peine d'étudier la valeur de cet item? Il est permis d'en douter.

15. Trouver des rimes.

Une autre variante du test d'associations est celui de trouver des rimes. Il semblerait bien à première vue qu'il ne faille pas beaucoup de capacités intellectuelles pour réussir ce test. Comme le fait remarquer Terman²³ des associations de sons peuvent paraître beaucoup plus simples que des associations de cause et effet, de partie au tout, d'opposés, etc. Cependant, lorsque les auteurs du test examinent les données recueillies sur ce test ils y trouvent une forte corrélation avec l'intelligence générale. Les enfants de neuf ans qui possèdent un développement mental de dix ans ou plus réussissent presque toujours à trouver les rimes tandis que les enfants de neuf ans qui atteignent un niveau de huit ans réussissent rarement. Un test qui dénote ainsi une haute corrélation avec l'échelle en général doit nécessairement être valide.

²³ Terman, op. cit., p. 249.

Pour trouver des rimes, il faut chercher des associations verbales sous la direction d'une idée maîtresse. Chaque mot suscite une foule d'associations et certaines viennent à l'idée plus facilement que d'autres. Celles qui n'ont aucun rapport avec le but qu'on se propose, doivent donc être éliminées. Il faut donc que l'idée maîtresse soit assez forte pour diriger la pensée vers le but proposé. Même les débiles mentaux ont un vocabulaire suffisant pour trouver une foule de mots qui riment avec 'jour' par exemple. Leur difficulté vient de leur incapacité de guider leur pensée sous la direction de l'idée maîtresse pour ne choisir que les mots qui riment. Le mot 'jour' pourra leur suggérer 'bonjour' et oubliant ensuite l'idée dirigeante, ils continueront avec des mots comme 'matin', 'soir', etc. Le manque d'associations cause aussi un grand nombre d'échecs. Terman fait encore remarquer²⁴ que les adultes font moins bonne figure que des enfants normaux qui ont le même niveau mental parce que les exercices scolaires donnent à leur esprit une plus grande souplesse qui se perd graduellement avec les préoccupations plus nombreuses de la vie.

Ce test est celui qui a causé le plus de controverse quant à sa position dans l'échelle. Binet le place à douze ans dans son échelle de 1908 pour ensuite le monter à quinze ans dans celle de 1911. Kuhlmann le garde à douze ans tandis que Goddard le descend à onze ans. Terman le propose à neuf ans et Amoss fait de même. Le désaccord vient de la façon différente de présenter et de coter le test aussi bien que du choix de mots. Binet ne donne qu'un essai rendant ainsi le test beaucoup plus

24 Loc. cit.

difficile que lorsqu'on permet trois essais et qu'on exige que deux réussites. De plus le mot utilisé par Binet (obéissance) est beaucoup plus difficile que les mots d'une syllabe employés par Terman (day, mill, spring). Les statistiques de Terman²⁵ montrent que dans quatre séries de ce genre, soixante-deux pour cent réussissent à neuf ans, quatre-vingt-un pour cent à dix ans, quatre-vingt-trois pour cent à onze ans et quatre-vingt-quatorze pour cent à douze ans. Dans le New Revised Stanford-Binet Terman emploie des mots nouveaux (date, head, cap) parce qu'il avait trouvé que les mots de sa révision précédente n'étaient pas satisfaisants.

Amoss présente ce test de la même façon que le fait Terman dans le Stanford-Binet. Il n'y fait que deux modifications dont la première est de substituer le mot 'bill' à 'hill', et la deuxième de dire au sujet: "Dis-moi aussi vite que tu peux tous les mots qui riment avec . . ." Voici comment Terman propose le test: "Je vais te donner un mot et tu auras une minute pour trouver autant de mots que tu peux qui riment avec . . ." Terman prévient de ne rien dire de plus au sujet qui pourrait lui suggérer qu'il lui faut se hâter parce que pour certains enfants nerveux cette idée les jette dans la confusion. Ces petits détails peuvent paraître sans importance à première vue, mais il ne faut pas travailler longtemps sur les tests pour se rendre compte qu'ils peuvent influencer facilement les résultats.

25 Terman, op. cit., p. 251.

16. Disposer des poids.

Disposer des poids selon leur pesanteur mesure l'intelligence pratique d'un sujet plutôt que sa capacité d'abstraction. C'est l'aspect performance qui est mesuré, c'est-à-dire l'habileté à faire quelque chose plutôt que l'habileté de s'exprimer par des paroles. Il jette de la lumière sur le jugement pratique du sujet, jugement qui joue un rôle important dans la vie quotidienne de chaque individu.

Binet plaça tout d'abord ce test à neuf ans mais il le descendit à huit ans dans sa révision de 1911. Il est retenu à neuf ans dans la plupart des révisions de l'échelle métrique. La position du test dans l'échelle dépend des poids qui sont utilisés, de la présentation et du cotage du test. Kuhlmann utilise des poids de trois, neuf, dix-huit, vingt-sept, trente-six et quarante-cinq grammes. Terman emploie des poids de trois, six, neuf, douze et quinze grammes. Le discernement est certainement plus facile à faire entre des poids qui diffèrent de neuf grammes qu'entre ceux qui ne diffèrent que de trois grammes. De plus, si on exige qu'une réussite sur trois le test peut être placé un ou deux ans plus bas que si on exige trois réussites sur trois essais.

Une grande partie de la difficulté du test vient de la façon de le présenter, si oui ou non on laisse trouver à l'enfant la façon de procéder. Kuhlmann et Bobertag illustrent la façon de procéder en exigeant que l'examineur soupèse tout d'abord les blocs et les arrange en ordre de pesanteur tandis que le sujet regarde faire. Terman préfère garder ce test dans sa

forme originale parce que, dit-il,²⁶ si on illustre on enlève au test les caractéristiques qui lui donnent sa valeur. Il ne devient alors qu'un test de discernement sensoriel et ce n'est pas là le but du test. On veut savoir tout d'abord si l'enfant comprend la tâche qu'on lui impose et le but à atteindre; ensuite, s'il peut choisir une méthode appropriée à la réalisation du but; et finalement s'il est capable de ne pas perdre de vue la fin à atteindre tandis qu'il fait les pas nécessaires pour y arriver. Si élémentaires que puissent paraître les procédés qui sont impliqués, ils représentent néanmoins le prototype de tout comportement bien déterminé. Quelle que soit la profession qu'on exerce, chacun à sa façon et avec l'outillage dont il dispose, doit se fixer un but et choisir les moyens pour réaliser la fin.

Dans cette expérience-ci le sujet peut manquer une des trois exigences du test ou même toutes les trois. Tout d'abord il peut ne pas comprendre les directives et alors être incapable de se fixer un but; ensuite, bien qu'il comprenne ce qu'on attend de lui, il peut adopter une méthode absurde pour accomplir la tâche; en dernier lieu il peut perdre de vue la fin à réaliser et se mettre à jouer avec les blocs.

Il est très révélateur d'écrire d'une façon brève et concise la performance du sujet plutôt que de noter simplement un succès ou un échec. La méthode idéale est de soupeser chaque poids avant d'essayer de les placer, ensuite faire un premier arrangement et enfin le corriger en comparant les poids individuellement. Règle générale, on s'aperçoit qu'il

26 Ferman, op. cit., p. 238.

n'y a que les enfants très intelligents de dix ans et plus qui penseront à vérifier leur première disposition en soupesant chaque poids une dernière fois.

Amoss offre ce test comme test supplémentaire à neuf ans. Sa présentation est la même que celle de Terman mais sa façon de coter diffère. Si le sujet réussit deux essais sur trois il obtient un crédit de deux mois et s'il ne réussit qu'un essai et que dans un autre essai il n'y a pas plus d'une erreur on lui accorde un mois de crédit. Encore ici, un sujet peut obtenir un mois de crédit sur le C.I.E. là où il n'obtiendrait rien sur le Stanford-Binet.

17. Les absurdités.

Un test très ingénieux et très utile dans la mesure de l'intelligence est sans contredit la découverte d'absurdités. Il est très peu influencé par la scolarité du sujet. Comme les questions de compréhension on peut l'appeler un test de jugement si l'on emploie ce mot dans son sens familier. Une personne stupide ne voit pas les contradictions les plus absurdes parce que son jugement est très peu développé. Les idées restent isolées de sorte qu'elles seront gobées sans protestation de l'esprit critique.

Cet esprit critique commence à se développer quand l'enfant est très jeune et il se fortifie avec la richesse grandissante des associations d'idées. Selon Terman²⁷, pour l'enfant d'intelligence moyenne, cet esprit

27 Terman, op. cit., p. 259.

critique n'est égal à la tâche qu'on lui présente ici que vers dix ans. Les enfants d'un niveau mental de huit ans ne découvrent jamais plus de deux ou trois des cinq absurdités tandis qu'à douze ans l'esprit critique est tellement développé que la plupart des items sont presque toujours réussis.

Binet plaça ce test à onze ans dans l'échelle de 1908, mais il le descendit à dix ans dans sa révision de 1911. Goddard et Kuhlmann le retiennent à onze ans bien que la plupart des statistiques, les leurs comprises, justifient sa place à dix ans²⁸. Kuhlmann le placera à dix ans dans sa révision de 1922. La plupart des auteurs exigent trois réussites sur cinq essais.

Terman trouve que les deuxième et cinquième absurdités de la liste de Binet présentent des difficultés marquées pour des enfants de dix ans (voir le Tableau VI); la cinquième est trop subtile et la deuxième présente une difficulté linguistique parce que cette expression est acceptée dans la langage courant de certains milieux. Pour remplacer ces deux absurdités il y substitue les numéros un et deux du Tableau VI. Terman considère les cinq phrases absurdes qu'il présente comme étant presque d'égale difficulté et en général plus faciles que celles de Binet²⁹. En effet, on peut constater en comparant les phrases de ces deux auteurs que Terman a rendu les siennes plus simples tout en gardant l'idée de Binet.

28 Loco cit.

29 Loco cit.

Tableau VI.- Les absurdités proposées par différents auteurs à dix ans.

Binet et Kuhlmann	Terman - 1916	Amoss
1° Un malheureux cycliste a eu la tête fracassée et il est mort sur le coup; on l'a transporté à l'hôpital et on craint bien qu'il ne puisse en réchapper.	1° Un homme disait: Je connais un chemin qui va en descendant tout le temps de ma maison jusqu'à la ville et en descendant tout le temps pour revenir de la ville à la maison.	1° Un homme dit: J'ai trois frères, Jean, Joseph et moi.
2° J'ai trois frères: Paul, Ernest et moi.	2° Un ingénieur disait que plus il y avait de wagons à son train, plus il pouvait aller vite.	2° Un ingénieur disait que plus il y avait de wagons de fret à son train plus il pouvait aller vite.
3° On a trouvé hier sur les fortifications le corps d'une malheureuse jeune fille coupé en dix-huit morceaux. On croit qu'elle s'est tuée elle-même.	3° Hier la police a trouvé le corps d'une jeune fille coupé en dix-huit morceaux. On croit qu'elle s'est tuée elle-même.	3° Hier la police a trouvé le corps d'un homme coupé en dix-huit morceaux et on craint qu'il se soit tué lui-même.
4° Il y a eu hier un accident de chemin de fer mais ça n'est pas grave; le nombre des morts est seulement de 48.	4° Il y a eu hier un accident de chemin de fer mais ça n'est pas grave; le nombre de morts est seulement de 48.	4° Un garçon dit: Je marche en montant tout le temps de ma maison à l'école, et je marche en montant tout le temps de l'école à la maison.
5° Quelqu'un disait: Si je me tue un jour de désespoir, ce ne sera pas un vendredi que je choisirai, car le vendredi est un mauvais jour qui me porterait malheur.	5° Un cycliste, jeté de sa bicyclette dans un accident, se frappa la tête contre une pierre et fut tué instantanément. On l'a ramassé et on l'a emporté à l'hôpital et on craint bien qu'il ne puisse en réchapper.	5° Un garçon faisant face à un vent froid d'hiver en s'en allant à l'école le matin dit: "J'espère que le vent va changer complètement et souffler dans la direction opposée avant que je retourne à la maison pour ne pas que j'aie à lui faire face.

En consultant le Tableau VI, nous voyons que dans la liste d'absurdités proposées par Amoss la première est la même que le numéro deux de Binet, les trois suivantes sont copiées de Terman (Amoss ne fait que changer un mot par-ci, par-là) et la dernière est sa contribution personnelle. Nous avons vu que la première phrase est inacceptable pour la raison citée plus haut par Terman. Nous ne connaissons pas la valeur de la dernière bien qu'elle semble satisfaisante. Il n'y aurait que l'étude des items particuliers qui renseignerait si oui ou non elle est bien placée à ce niveau-ci.

Pour quatre réussites, Amoss accorde les deux mois de crédit comme le fait Terman mais il accorde aussi un mois de crédit pour trois réussites. Voici encore un test sur lequel un sujet peut obtenir un mois de crédit de plus sur le C.I.E. que sur le Stanford-Binet. Terman présente de nombreux exemples de réponses satisfaisantes et de réponses qui ne le sont pas, rendant ainsi la correction de cet item aussi objective que possible. Amoss omet ce point et il arrivera maintes fois que l'examinateur hésitera à classer certaines réponses comme des échecs ou des réussites.

18. Nommer des mots en une minute.

Bien que ce test souffre de certaines imperfections, il a sa valeur parce qu'il renseigne sur la richesse et la variété d'associations faites au préalable avec des mots ordinaires et qu'il permet d'observer la facilité avec laquelle elles peuvent se rétablir. Le jeune sujet ou le sujet arriéré ne trouvera qu'un mot pour chaque idée tandis que le sujet plus intelligent énumérera une demi-douzaine de mots pour chaque idée qui lui

traversera l'esprit. Ce dernier accomplit plus avec moins d'effort, mais ce procédé requiert de l'intelligence et de la volonté pour ne pas perdre de temps avec des mots isolés.

On est souvent étonné de la pauvreté d'associations qu'un petit test révèle chez les sujets retardés. Il arrive souvent qu'ils ne peuvent trouver un seul mot durant une vingtaine de secondes ou plus. Il serait intéressant alors de pouvoir pénétrer la conscience du sujet. Il y a certainement quelque chose, mais les idées sont trop vagues et trop nébuleuses pour se cristalliser en mots.

Il est très instructif de noter le genre de réponses données. Certains sujets, surtout ceux qui ont un niveau mental de huit ou neuf ans donnent des mots isolés. Le mot 'chapeau', par exemple, ne leur suggère rien d'autre et ils passent à un autre mot sans se rendre compte que les chapeaux diffèrent en forme, en couleur, qu'ils ont différents usages et différents accessoires et qu'en énumérant ceux-ci, ils pourraient trouver une foule de mots. D'autres sujets prennent vite avantage de telles relations et nomment plusieurs parties d'un objet ou plusieurs objets appartenant à la même classe. Ainsi, le mot 'crayon' suggère: papier, plume, encre, encrier, cahier, livre, pupitre, chaise, classe, école, etc. Un troisième genre de réponse consiste à nommer les objets dans la salle. Il est dommage que ceci arrive parce que les salles où les tests sont donnés varient beaucoup dans leur ameublement. Le sujet qui choisit cette méthode est handicapé si la pièce est plutôt nue. Un quatrième genre de réponse est celui qui contient une forte proportion de mots abstraits ou hors de l'ordinaire. C'est évidemment un genre de réponse supérieur qu'on

trouve plutôt rarement et seulement chez des sujets qui ont un niveau intellectuel de plus de onze ans.

Il arrive parfois qu'un enfant avec un niveau intellectuel de onze ou douze ans fasse mauvaise figure sur ce test. La chose arrive parce qu'il persiste à épuiser tous les mots d'une classe avant de passer à une nouvelle série ou parce qu'il perd son temps à chercher de beaux grands mots. Ceci explique sans doute pourquoi certains adultes font si pauvre figure. Ils sacrifient la vitesse à la logique des idées qu'ils veulent poursuivre à tout prix.

Terman ne croit pas que ce défaut soit assez sérieux pour éliminer le test. Dans la majorité des cas il trouve qu'il y a corrélation entre le résultat obtenu sur ce test et le résultat global. Les sujets de plus d'un ou deux ans au-dessous de l'âge mental de dix ans réussissent rarement tandis que les sujets d'un ou deux ans au-dessus de dix ans faillissent rarement. Terman considère de plus que le test a une signification particulière selon le genre de mots donnés. L'énumération de mots abstraits ou hors de l'ordinaire indiquerait une intelligence supérieure tandis que les répétitions des mêmes mots dénoteraient cette tendance à stéréotyper qui est si fréquente chez les débiles³⁰.

Binet proposa tout d'abord ce test à onze ans mais il le monta à douze ans dans sa révision de 1911. Goddard et Kuhlmann, selon Terman³¹ le gardent à onze ans bien que les statistiques de Goddard montrent qu'il

³⁰ Terman, op. cit., p. 275.

³¹ Loco cit.

serait bien placé à dix ans. Toutes les statistiques américaines s'accordent à démontrer que le test est assez facile pour dix ans.

Dans le Stanford-Binet Terman exige que soixante mots soient nommés dans l'espace de trois minutes. Amoss en exige trente en une minute pour obtenir deux mois de crédit et vingt-quatre pour un mois de crédit. Dans la formule L du New Revised Stanford-Binet Terman exigera vingt-huit mots en une minute. Amoss se montre donc un peu plus sévère que Terman, mais par contre le sujet obtient un mois de crédit pour vingt-quatre mots. Il est évident que les résultats obtenus sur les deux tests ne deviennent plus directement comparables. Amoss peut-il justifier ces crédits partiels? Peut-être, mais il faudra qu'il le prouve par sa standardisation.

19. Le plan de recherche.

On peut fort bien appeler cet item un test de jugement pratique. Contrairement à la plupart des tests celui-ci fournit l'occasion au sujet de montrer comment il peut faire face à une situation réelle plutôt qu'à une situation imaginée. Des tests comme celui-ci appuient davantage sur le concret et contrebalance la trop grande importance donnée au raisonnement abstrait et à la compréhension du langage. En effet, une échelle de tests verbaux ne donne pas une image complète de l'intelligence d'un sujet puisqu'elle ne mesure pas l'aspect performance. L'idéal semblerait bien être un test comme le Wechsler-Bellevue où il y a autant de tests de performance que de tests verbaux. Si on accepte comme définition de l'intelligence en psychologie comme étant la capacité de faire face à de nouveaux problèmes d'une façon rationnelle, il faut bien se rendre à l'évidence qu'il y a

autant de problèmes concrets à résoudre que d'abstraites. Un grand nombre de sujets qui seraient classés comme débiles sur un test verbal font suffisamment preuve d'intelligence pratique sur un test de performance pour ne pas être classés dans cette catégorie puisqu'ils montrent qu'ils peuvent se tirer d'affaire assez bien.

Terman présente ce test à deux niveaux différents dans le Stanford-Binet, tout d'abord à huit ans et ensuite à douze ans. Il est très indulgent à huit ans et il cote le test libéralement, mais à douze ans il faut que la reproduction remplisse toutes les conditions logiques du problème et pour assurer l'uniformité dans la façon de coter il donne plusieurs illustrations de réponses acceptables et de celles qui ne le sont pas. Il présente ce test à treize ans dans la formule L du New Revised Stanford-Binet sans doute parce qu'il constata que le pourcentage de réussites à douze ans n'était pas suffisant.

Amoss présente ce test à douze ans. La présentation du test est la même que celle de Terman à l'exception qu'il substitue un dollar à la balle perdue dans le champ. Ces directives pour coter le test ne sont pas claires:

If the candidate indicates a method of search, (1) which will cover the field with approximate uniformity, (2) so that no path crosses any other path save perhaps at the entrance, (3) and so that some one diameter of the circle will intersect at least five pathways, credit three months.³²

Ce qu'Amoss veut dire par (3) est bien difficile à déchiffrer et manque beaucoup de clarté. Il aurait assuré beaucoup plus d'uniformité si, comme Terman, il avait illustré avec des dessins. Il est certain qu'à moins que les examinateurs connaissent le Stanford-Binet et s'y réfèrent pour coter ce test-ci, il y aura une grande divergence dans la façon de coter.

³² Amoss, op. cit., p. 59.

Tous ces détails marquent bien la différence entre une échelle bien construite et une qui ne l'est pas. Amoss montre bien peu de souci pour le fini et le détail.

20. Les phrases en désordre.

Les phrases en désordre remplissent la condition principale d'un bon test, c'est-à-dire que le pourcentage de réussites augmente rapidement et régulièrement à mesure que l'âge mental augmente. De plus, le succès dans ce test dépend très peu de l'expérience acquise par le sujet.

On peut considérer ce test comme une variante du completion test d'Ebbinghaus. En effet, Binet nous dit que ce test lui a été suggéré par celui d'Ebbinghaus. Ce dernier omet des parties de phrases et il demande au sujet de suppléer aux omissions tandis que Binet présente toutes les parties en désordre et exige qu'on fasse une bonne phrase en mettant tous les mots en ordre. Du point de vue psychologique, les deux tests se ressemblent parce qu'il faut que le sujet réunisse des fragments pour en faire un tout qui a du sens.

On trouve tout d'abord ce test à onze ans dans l'échelle métrique de 1908 et à douze ans dans celle de 1911. Goddard et Kuhlmann le retiennent à onze ans mais toutes les statistiques excepté celles de Goddard démontrent que le test est bien placé à douze ans. Ces statistiques s'accordent très bien comme le prouve le pourcentage de réussites à douze ans dans les recherches faites par divers auteurs: Binet, 66; Kuhlmann, 68; Bobertag, 78; Dougherty, 64; Strong, 72; les séries Stanford (1911), 70;

les séries Stanford (1913), 57; les séries Stanford (1914), 62³³.

Amoss comme Terman présente ce test à douze ans. Les directives sont les mêmes. Amoss présente quatre phrases tandis que Terman n'en présente que trois. Ce dernier n'exige que deux réussites sur trois essais pour accorder le crédit total tandis qu'Amoss exige trois réussites sur quatre pour accorder le crédit total de trois mois, deux réussites pour deux mois de crédit et une réussite pour un mois de crédit. Ici encore se présente le problème des crédits partiels qui empêchent les deux tests d'être directement comparables.

De plus, Amoss présente des phrases tout à fait différentes de celles qui sont proposées par Terman. Ce dernier propose les trois phrases suivantes:

1. FOR THE STARTED AN WE COUNTRY EARLY AT HOUR.
2. TO ASKED PAPER BY TEACHER CORRECT I BY.
3. A DEFENDS DOG GOOD HIS BRAVELY MASTER.

tandis qu'Amoss propose celles-ci:

1. APPLE TASTED UP AND HE IT PICKED AN.
2. TO ASKED COAT MOTHER MY MEND I BY.
3. GOOD IS TO /RONG SHE TOO ANYTHING DO.
4. WOULD IT HE KEEP THOUGHT TO LIKE HE.

Nous pouvons constater à première vue que la seconde phrase d'Amoss imite d'assez près la seconde phrase proposée par Terman. Il est évident que les deux dernières phrases et surtout la dernière, proposées par Amoss sont beaucoup plus difficiles que celles de Terman. Pour connaître si ces phrases sont valides pour douze ans il faudrait que le travail de standardisation soit fait et nous doutons fort qu'il l'ait été puisqu'on ne trouve

33 Terman, op. cit., p. 290.

rien d'écrit à ce sujet.

Dans le New Revised Stanford-Binet Terman propose les mêmes phrases à treize ans sans doute parce que sa nouvelle standardisation à montrer qu'elles étaient trop difficiles pour douze ans. Il est donc permis de conclure que les phrases utilisées par Amoss qui semblent plus difficiles que celles de Terman devraient aussi être placées à treize ans sinon plus haut dans l'échelle. Voilà donc encore un item de validité douteuse que nous pouvons ajouter à notre liste déjà longue.

21. Le problème d'induction.

Ce test imaginé par Terman consiste à plier une feuille de papier en deux, à découper un petit morceau au milieu du côté plié et demander au sujet combien il y aura de trous une fois la feuille dépliée. On répète le procédé six fois en se servant de feuilles différentes chaque fois. On ajoute un pli nouveau chaque fois et on ne fait qu'une seule découpe. Après chaque découpe on demande au sujet combien il y a de trous. Quelle que soit la réponse, on doit déplier le papier pour que le sujet soit à même de constater si sa réponse est bonne ou mauvaise. Le nombre de trous est double à chaque nouveau pli ajouté. Il faut qu'au sixième essai le sujet puisse répondre qu'il y aura trente-deux trous et qu'il puisse donner la règle, c'est-à-dire que le nombre de trous est doublé à chaque fois qu'on ajoute un pli. Le test est réussi seulement quand le sujet peut donner la règle, et on ne lui accorderait pas les points s'il donnait la bonne réponse sans pouvoir expliquer le principe. Règle générale, plus le niveau intellectuel du sujet est élevé moins il

faut donner de démonstrations pour que le sujet saisisse la règle.

Terman utilisa ce test pour la première fois dans une étude comparative de garçons brillants et de garçons bornés, mais il ne le standardisa qu'en 1914. Ses données nombreuses indiqueraient que c'est un bon test d'intelligence: cinquante-neuf pour-cent des écoliers de quatorze ans avec un quotient intellectuel entre 96 et 105 réussirent le test; quarante et un pour-cent des écoliers de quatorze ans avec un quotient intellectuel au-dessous de 96 réussirent; soixante et onze pour-cent des sujets avec un quotient au-dessus de 105 réussirent; seulement dix pour-cent des adultes moyens et moins de cinq pour-cent d'adultes manquèrent le test³⁴. Le test s'accorde donc bien avec les résultats obtenus sur l'échelle tout entière. Le test a encore l'avantage d'être très peu influencé par la scolarité et à l'exception des différences intellectuelles, il est peu influencé par l'âge. De plus, il suscite beaucoup d'intérêt chez le sujet.

Terman a appelé ce test un "test d'induction". On pourrait supposer qu'on arrive à la solution par déduction ou par une analyse logique a priori du principe impliqué, mais ceci arrive rarement. Ordinairement ce n'est qu'après une ou deux erreurs qui sont montrées au sujet en dépliant la feuille, qu'il saisira le principe.

Terman emploie à quatorze ans dans le Stanford-Binet et dans le New Revised Stanford-Binet. Pour obtenir les quatre mois de crédit accordés aux tests de cet âge, il faut que le sujet puisse donner la règle

³⁴ Terman, op. cit., p. 312.

sans cela il n'obtient aucun crédit. Amoss présente ce test de la même façon que Terman mais il le cote différemment. Il accorde deux mois de crédit pour la bonne réponse sans la règle, et deux mois pour la règle. Cette façon de coter est beaucoup moins sévère que celle de Terman et elle diminue beaucoup la difficulté du test. Cependant Amoss place son test au même niveau que celui de Terman. Encore une fois un sujet pourra obtenir deux mois de crédit de plus sur le C.I.E. que sur le Stanford-Binet. Il semble bien qu'Amoss fait ces changements pour l'unique raison de ne pas être accusé d'avoir tout simplement reproduit le Stanford-Binet.

22. Renverser les aiguilles de l'horloge.

Dans ce test-ci l'examineur indique une heure au sujet et il lui demande s'il peut se représenter où seraient la grande aiguille et la petite aiguille à l'heure dite. Ensuite il lui demande quelle heure il serait si la grande aiguille prenait la place de la petite aiguille et si la petite aiguille prenait celle de la grande. Il faut que l'examineur s'assure que le sujet ne consulte ni une horloge, ni une montre, et qu'il ne s'aide pas en faisant un dessin.

Il semblerait que le succès sur ce test dépend en grande partie de la capacité d'évoquer des images visuelles. Ce manque de capacité explique peut-être le pourcentage élevé d'échecs chez les adultes. La scolarité semble n'avoir aucune influence sur le pourcentage des réussites. Par contre, c'est un des tests qui peut être le plus influencé par une connaissance acquise de sujets qui ont déjà subi le test. C'est pourquoi, nous

dit Terman³⁵, il faudrait essayer d'autres positions des aiguilles de l'horloge qui seraient d'égale difficulté et qu'on pourrait substituer à celles qui sont proposées. Terman n'a pas fait ce travail et on ne doit pas en utiliser d'autres parce que, pour une raison inconnue, nous dit-il encore, certaines positions sont plus difficiles les unes que les autres. Il a constaté par expérience que 2.46 est beaucoup plus difficile que 8.10 et 6.22. Cette première position donne deux fois plus d'échecs que les deux dernières³⁶.

Terman supprimera ce test dans le New Revised Stanford-Binet parce qu'il ne le trouve pas satisfaisant pour les raisons mentionnées plus haut.

Ce test utilisé par Binet dans son échelle de 1905 fut supprimé dans les révisions de 1908 et de 1911. Goddard et Kuhlmann le conservent dans leur révision mais ils le placent à quinze ans. Ils ne permettent que deux essais qui doivent être réussis pour obtenir le crédit. Ces deux auteurs n'indiquent pas la marge d'erreur permise.

Les trois positions des aiguilles que Terman propose sont 6.22, 8.10 et 2.46. La marge d'erreur permise va de 4.30 à 4.35, 1.40 à 1.45 et 9.10 à 9.15 respectivement.

Amoss propose les quatre positions suivantes: 6.22, 8.08, 4.57 et 2.46. Les réponses sont acceptables si elles sont de 4.30 à 4.34, 1.40 à 1.44, 11.21 à 11.25 et 9.11 à 9.15 respectivement.

35 Terman, op. cit., p. 322.

36 Loco cit.

En comparant les heures proposées par Terman et par Amoss on peut constater que la première et la dernière sont les mêmes; la troisième est l'invention d'Amoss et la deuxième varie de deux minutes. Nous sommes portés à croire qu'Amoss a rendu cette troisième position plus difficile car il est plus facile d'évoquer la position des aiguilles à une des divisions du cadran comme 8.05 ou 8.10 que 8.08. Quel est le degré de difficulté de la position 4.57? Nous n'en savons rien. Comme la position 2.46 qui est beaucoup plus difficile que les deux premières, il se peut fort bien que 4.57 le soit aussi. Il n'y aurait que l'expérience qui pourrait nous renseigner sur ce point. Amoss s'est-il donné la peine d'en chercher le degré de difficulté? Il est permis d'en douter.

Comme nous le soulignons plus haut Terman exige deux réussites sur trois essais pour accorder les deux mois de crédit. Amoss augmente énormément la difficulté de ce test en accordant un mois de crédit pour chaque essai répondu correctement. Pour obtenir le crédit total il faudrait que le sujet réussisse les quatre essais. Kuhlmann et Goddard exigent deux bonnes réponses sur deux mais ils placent le test à quinze ans. Comme le propose Amoss ce test serait probablement bien placé à adulte moyen et peut-être même à adulte supérieur.

23. Les problèmes d'ingéniosité.

Ce test a été appelé un test d'ingéniosité car, en effet, il faut être ingénieux pour résoudre les problèmes qui sont présentés ici.

Voici comment on présente les problèmes au sujet: "Une mère envoie son fils à la rivière pour qu'il rapporte exactement sept copines d'eau.

Elle lui donne pour cela une chaudière de trois chopines et une chaudière de cinq chopines. Comment le garçon pourra-t-il rapporter exactement sept chopines en ne se servant que des deux chopines sans deviner? Vous devez commencer par remplir la chaudière de cinq chopines. Remarquez bien, avec une chaudière de trois chopines et une chaudière de cinq chopines vous devez rapporter exactement sept chopines. Le second problème se pose comme le premier, mais avec une chaudière de cinq chopines et une chaudière de sept chopines il faut rapporter exactement huit chopines en commençant par remplir la chaudière de cinq chopines. Dans le troisième problème, avec une chaudière de quatre chopines et une chaudière de neuf chopines il faut rapporter sept chopines en commençant par remplir la chaudière de quatre chopines.

Devant ces problèmes le sujet fait face à une difficulté qu'il lui faut résoudre. Il a besoin de jugement pratique aussi bien qu'un certain degré d'ingéniosité. Il doit examiner diverses possibilités qu'il rejette ou accepte. Ce don d'invention fait partie sous une forme ou une autre de toute activité mentale. Il dépend surtout d'une imagination constructive c'est-à-dire de la capacité d'abstraire quelque chose de ce qui est immédiatement présent aux sens et d'évoquer les nouvelles situations avec leurs possibilités et leurs conséquences. Les images sont réunies pour faire de nouvelles combinaisons.

Les différences individuelles marquées ne dépendent pas tellement du discernement sensible mais beaucoup plus de la projection juste des éléments sensoriels dans la mémoire et l'imagination. C'est là une capacité que tout être humain possède à un degré plus ou moins marqué. Tout

test qui nous donne une connaissance de la capacité d'un sujet dans ce domaine a sa valeur.

Ce test, comme le problème d'induction, fut imaginé par Terman et utilisé pour la première fois en 1905 dans une étude comparative de garçons brillants et de garçons arriérés, mais il ne fut pas standardisé à ce moment. Terman trouva par la suite qu'il appartenait à un niveau mental beaucoup plus élevé qu'il ne l'avait cru tout d'abord. Selon ses statistiques³⁷ il n'y a qu'un très petit nombre de sujets avec un âge mental au-dessous de quatorze ans qui passent le test, à peu près les deux tiers des sujets d'intelligence moyenne le manquent et soixante-quinze pour-cent des adultes supérieurs le réussissent. C'est un test qui est très peu influencé par la scolarité puisque des hommes d'affaires avec peu de scolarité réussissent un peu mieux que des étudiants d'école secondaire.

Si le sujet ne trouve pas la solution du premier problème, l'examineur lui explique et lui présente ensuite le second problème. Il fait de même pour le troisième. Dans le New Revised Stanford-Binet Terman présente les problèmes d'ingéniosité à quatorze ans mais il rend le premier problème assez facile pour que le sujet le trouve sans difficulté. Si le sujet ne trouve pas la solution, on ne lui explique pas et on ne lui présente pas les deux autres.

Amoss copie textuellement ses problèmes d'ingéniosité du Stanford-Binet. Il n'y a que sa façon de coter qui diffère. Terman n'exige que deux réussites pour accorder le crédit total tandis qu'Amoss accorde deux

37 Terman, op. cit., p. 348.

mois de crédit pour chaque problème. Il faut donc trois réussites pour obtenir le crédit total sur le C.I.E. Cette façon de coter augmente beaucoup la difficulté du test. Les sujets assez intelligents pour réussir deux items mais non pas trois sont placés à un désavantage puisqu'ils n'obtiennent que quatre mois de crédit sur le C.I.E. là où ils en obtiendraient six sur le Stanford-Binet. Par contre, les sujets moins intelligents qui ne réussissent qu'un item sont favorisés puisqu'ils peuvent obtenir deux mois de crédit là où ils n'en obtiendraient aucun sur le Stanford-Binet.

24. La compréhension de la lecture.

Ce test a dû être imaginé par Amoss car nous n'avons pas pu en trouver la source. Il consiste à présenter le texte suivant:

The group of eight circular units . . . is a section selected by chance from an extensive surface. Without introducing any additional or different shaped units into your drawing reconstruct the group to illustrate change from an indefinite incoherent homogeneity to a definite coherent heterogeneity. You will have two minutes.³⁸

Ce test n'est pas tellement un test de compréhension de la lecture d'un texte mais plutôt un test de vocabulaire, car si le sujet ne comprend pas les mots 'homogénéité' 'hétérogénéité' 'cohérent' et 'incohérent' il ne pourra pas savoir ce qu'il doit faire. Ce test sera donc influencé par la scolarité du sujet et un sujet illettré y fera bien mauvaise figure car ses chances seront minces de comprendre un vocabulaire aussi compliqué.

³⁸ Amoss, op. cit., p. 28.

25. Le dessin au miroir.

Ce test présenté par Amoss à adulte supérieur se retrouve nulle part ailleurs dans une échelle pour mesurer l'intelligence. C'est un test d'associations et d'apprentissage plutôt qu'un test d'intelligence. En effet, ce test oblige le sujet à former une nouvelle série d'associations qui sont opposées à des associations stéréotypées par plusieurs années d'expérience quotidienne.

En traçant un dessin ordinaire les mouvements de la main sont guidés par la perception visuelle du dessin et par les sensations du toucher créées par le mouvement du crayon. Si le dessin est vu dans un miroir, les relations naturelles sont renversées de sorte qu'une nouvelle série d'associations doivent être établies entre l'oeil et la main. La rapidité et la facilité avec laquelle ces nouvelles associations sont établies peuvent être considérées comme un index d'apprentissage.

Dans le test proposé par Amoss l'examineur montre au sujet la carte sur laquelle est dessiné un carré traversé de deux diagonales. Il place ensuite un petit miroir de trois pouces par quatre pouces dans une position verticale en face du sujet devant le livret d'examen ouvert à une page blanche. Il donne un crayon au sujet, et place ensuite un écran en carton d'au moins huit pouces par douze pouces au-dessus du livret d'examen de façon que le sujet puisse voir la réflexion de la page blanche dans le miroir mais qu'il ne puisse pas la voir directement. Il lui dit ensuite: "Je veux que vous dessiniez dans cet espace le carré avec les deux diagonales tel que vous le voyez sur cette carte. Vous devez fixer les yeux

sur le miroir et guider vos efforts par la réflexion que vous y voyez. Vous pouvez essayer autant de fois que vous le voulez mais vous n'aurez que trois minutes pour compléter le dessin."

Amoss attribue à ce test une valeur de six mois répartis ainsi: un mois pour un rectangle dont le côté le plus long n'est pas plus de 1 fois $1/4$ le côté le plus court et dont les angles ne varient pas plus de vingt degrés d'un angle droit; deux mois pour la première diagonale et trois mois pour la deuxième diagonale.

En fait, que mesure ce test tel que proposé par Amoss? Pour répondre à cette question, il faudrait faire une série d'expériences qui constitueraient en elles-mêmes toute une recherche. Nous citerons donc les résultats de recherches qui ont déjà été faites sur les dessins au miroir et leur rapport avec l'intelligence:

Burt reports a correlation between speed and estimated intelligence of 0.67, P.E. .07, for elementary school boys; and of 0.54, P.E. .14, for preparatory school boys. In another group of English school children a correlation of 0.60 was found according to Burt and Moore. Miss Calfee, however, found no such relations in her group of elementary school children chosen to duplicate Burt's conditions: here the correlation with school grades was virtually zero (0.07); similarly, in the college students the correlation with grades was -.07 in the case of the men and 0.19 in the case of the women. The author was able to discern no constant differences between the work of five dull and five bright boys.⁴¹

Dans le test d'Amoss, il n'est pas question de mesurer la vitesse avec laquelle le sujet acquiert de nouvelles associations ni le nombre d'essais ou d'erreurs qu'il fait, mais tout simplement de savoir s'il

41 Guy Montrose Whipple, Manual of Mental and Physical Tests, Part II: Complex Processes, Baltimore, Warwick & York Inc., 1921, p. 494.

peut dessiner le carré et les deux diagonales dans un temps maximum de trois minutes. Comme le rapporte Burt, s'il y a une corrélation de 0,67 entre la vitesse d'acquisition et l'intelligence ceci ne veut pas dire que la même corrélation existe entre le test proposé par Amoss et l'intelligence. Tel que proposé par Amoss nous doutons fort que ce test mesure l'intelligence et nous sommes portés à croire qu'il n'a rien à faire dans une échelle qui mesure l'intelligence générale.

Dans ce troisième chapitre nous avons étudié les vingt-cinq items particuliers du C.I.E. qui n'apparaissent qu'à un endroit, et comme pour les séries d'items nous les avons comparés avec ceux de Binet, de Kuhlmann et de Terman pour ensuite en faire une critique. Nous avons vu que vingt-trois des vingt-cinq items ont été pris presque textuellement du Stanford-Binet. Amoss ne fait que changer quelques mots ici et là. Le changement le plus important est dans sa façon de coter; il accorde des crédits partiels dans quinze items là où Terman n'accorde aucun crédit. Il rend deux autres tests très difficiles en exigeant que tous les items soient réussis pour accorder le crédit total. Jusqu'à quel point ces changements influencent-ils le résultat global du test? Amoss a-t-il étudié la valeur de chaque item pour se rendre compte jusqu'à quel point il en augmente ou en diminue la difficulté et pour savoir si les tests sont bien placés au niveau où il les propose. C'est peu probable. De quel critère s'est-il servi pour justifier la position de tel ou tel item à tel endroit? Autant de questions qui demeurent sans réponse et qui font douter fortement de la valeur des résultats obtenus.

CHAPITRE IV

ETUDE CRITIQUE ET COMPARATIVE DU TEST EN GENERAL

Après avoir fait une étude critique et comparative des séries d'items et des items particuliers dans les deux chapitres précédents, nous ferons maintenant une étude critique et comparative du test en général. Cette étude portera tout d'abord sur la présentation du C.I.E.: allure générale, matériel et méthodologie; ensuite, puisqu'il s'agit d'un instrument psychométrique, nous en étudierons la constance, la validité et la technique d'administration. C'est surtout sur les qualités métrologiques que nous pourrons poser un jugement fondé sur la valeur de l'instrument qu'Amoss propose.

Dans son avant-propos Amoss nous dit qu'il craint de voir disparaître du marché le Stanford-Binet parce que Terman présente une nouvelle révision, le New Revised Stanford-Binet. Amoss trouve cette nouvelle révision trop dispendieuse et le matériel trop volumineux pour être pratique. Alors il propose un test peu dispendieux, peu volumineux, très simple selon ses prétentions, un test que les instituteurs peuvent apprendre à donner après avoir pris un cours assez bref.

En effet, son matériel se résume à peu de choses: un manuel de soixante-neuf pages, une série de cartes et quelques objets. Les objets nécessaires: une montre, un canif, un miroir et des poids, ne sont pas fournis avec le test de sorte que ce matériel n'étant pas standardisé sera choisi par chaque examinateur selon sa fantaisie. Il est certain

que ce matériel ne sera pas uniforme. Amoss semble juger ce point de peu d'importance; par contre, Terman présente toute une série d'objets standardisés qu'il ne permet pas d'être substitués par des objets semblables. Ce dernier a le souci du détail et il comprend la nécessité de l'uniformité dans un travail vraiment scientifique. Quand un test est standardisé dans certaines conditions, il faut que les expériences subséquentes se fassent dans des conditions identiques car autrement les résultats ne sont plus directement comparables. Amoss a-t-il compris cette nécessité? Il semble que non puisqu'il attache si peu d'importance au matériel de son test.

Son manuel contient soixante-neuf pages. Sa préface comprend trois pages et demie; vingt et une pages sont consacrées à l'illustration du matériel du test, c'est-à-dire aux reproductions d'images, aux dessins à reproduire et aux divers problèmes; deux pages et demie résument toutes les directives générales sur l'emploi du test; le reste du volume est consacré aux directives des tests aux différents âges. Nous ne trouvons rien sur la nature de son test, sur le développement et la standardisation de son échelle, sur l'analyse statistique de ses résultats. Cependant quand nous consultons les manuels de Terman, de Kuhlmann, de Wechsler, aussi bien que de tous les auteurs sérieux nous trouvons trois ou quatre chapitres consacrés à l'explication de la construction et de la standardisation de leurs échelles, aux directives générales et à l'interprétation des résultats obtenus. Amoss juge bon de résumer ses directives dans deux pages et demie! Il ne dit rien de la standardisation de son échelle.

C'est là une lacune inexcusable et nous croyons être justifiés en pensant que ce travail n'a pas été fait. Le bon Dr Amoss n'a pas l'esprit scientifique et ce n'est pas dans un bouquin comme le C.I.E. qu'on peut en apprendre long sur la technique des tests d'intelligence. La personne qui prétendrait savoir donner et interpréter un test d'intelligence après avoir étudié le C.I.E. aurait bien des prétentions et elle ferait preuve d'une belle naïveté.

Autres détails, ceux-ci au point de vue de la méthodologie: on ne trouve pas une seule référence au bas des pages, ni aucune bibliographie dans le manuel d'Amoss. Cependant nous savons que la plupart des tests sont copiés du Stanford-Binet. Par probité et surtout par souci scientifique un bon auteur fait connaître ses sources car autrement il fait sien le matériel d'un autre. Il ne faut pas aller bien loin dans l'étude du C.I.E. pour se rendre compte qu'Amoss se pare des livrées somptueuses de Terman mais que malheureusement pour lui il n'est pas de taille à les porter.

Puisqu'Amoss a imité Terman de si près nous exposerons les grandes lignes du développement et de la standardisation du New Revised Stanford-Binet¹ pour deux raisons: 1^o parce que c'est un des meilleurs tests individuels qui a été mis sur le marché, et 2^o pour bien faire comprendre quelle tâche formidable est celle de la construction d'un bon test individuel. Nous verrons ensuite en comparant le travail d'Amoss avec celui

1 Les renseignements qui vont suivre sur la présentation et la standardisation du New Revised Stanford-Binet ont été tirés du deuxième chapitre de Measuring Intelligence.

de Terman ce que le premier a fait ou plutôt ce qu'il n'a pas fait.

Un bon auteur renseigne toujours ses lecteurs sur le développement et la standardisation de son échelle. C'est ce qu'a fait Terman dans les trois premiers chapitres de Measuring Intelligence. Son premier pas fut la sélection des items qui devaient faire partie de son échelle. Parmi des milliers d'items qu'il pouvait choisir il donna la préférence au genre d'items qui fournissaient par expérience la plus haute corrélation avec un critère acceptable d'intelligence. Dans son choix il tint compte de certaines considérations pratiques comme la facilité de coter les items, l'intérêt que les tests suscitent chez les sujets, le temps requis pour faire passer les tests et la facilité d'administration. Il rejeta un bon nombre de tests excellents pour l'une ou l'autre des raisons énumérées plus haut.

Une fois que les items les plus prometteurs furent réunis, il en fit un essai expérimental. Cet essai n'était que préliminaire et avait pour but de faire l'élimination des tests les moins satisfaisants et de déterminer de façon approximative l'âge auquel les tests seraient employés. Cette expérience fut faite dans les environs de l'Université Stanford sur mille sujets dont les âges mentaux avaient été déterminés au préalable par l'administration du Stanford-Binet. Terman employa cette technique pour économiser du temps et il est entendu que sa sélection finale des tests et leur standardisation furent faites sur des populations non sélectionnées. Les items étaient trop nombreux pour en permettre l'administration à tout le groupe. Chaque test fut donc donné à une vingtaine

de sujets de chaque âge mental dans l'étendue où le test était applicable. Des courbes de pourcentage des sujets qui réussissaient ou échouaient aux différents âges mentaux furent dressées et la pente de ces courbes donnait une indication graphique de la validité des tests. La plupart des courbes couvraient toute l'étendue des groupes d'âges qui fournissaient cent pourcent d'échecs jusqu'à ceux qui fournissaient cent pourcent de réussites. Dans la préparation de l'essai final chaque test fut placé provisoirement à l'âge où cinquante pourcent des sujets réussissaient.

Après ce premier choix d'items, Terman en fit deux échelles provisoires pour un essai final. Il mit dans chacune un nombre suffisant de tests pour qu'il en reste assez pour deux échelles complètes une fois que les tests non satisfaisants seraient éliminés. Ces deux échelles furent administrées à plus de trois mille sujets sur lesquels fut basée la standardisation finale. La moitié des sujets passèrent la formule L avant la formule M et l'autre moitié passa la formule M avant la formule L. Tous les tests furent donnés par sept examinateurs à qui Terman avait enseigné pendant deux mois la technique d'administration de ses échelles. Il voulait s'assurer ainsi l'uniformité dans le procédé. Toutes les réponses devaient être écrites pour permettre la vérification.

Terman porta une attention particulière dans le choix de son échantillon pour qu'il soit tout à fait représentatif de la population. Tous ses sujets étaient nés aux Etats-Unis et de race blanche sans distinction de nationalité. Il choisit ses trois mille sujets dans dix-sept centres différents d'onze états. Il détermina le niveau socio-économique

des sujets d'après l'occupation du père. Il rangea les occupations dans cinq catégories différentes: le monde professionnel, le monde semi-professionnel, les métiers spécialisés, les métiers semi-spécialisés et les métiers non spécialisés. Il choisit le nombre de sujets de chaque milieu selon le pourcentage de parents dans chacune de ces cinq catégories. Il prit aussi un soin particulier du choix des sujets qui n'étaient pas encore d'âge scolaire.

Pour obtenir les deux échelles finales il lui fallut tout d'abord éliminer les tests qui s'étaient montrés les moins satisfaisants et arriver à un arrangement de tests pour que l'âge mental moyen de chaque groupe de sujets soit identique à l'âge moyen chronologique de ce groupe et qu'ainsi le quotient intellectuel moyen soit aussi près que possible de cent. Avant de faire ceci, toutes les formules d'examens furent vérifiées et cotées de nouveau par deux examinateurs. Pour assurer la précision dans les calculs statistiques les machines Hollerith furent employées. Plus de trente mille cartes sur lesquelles tous les renseignements possibles étaient compris, furent utilisées. Il était possible avec ces données de tracer des courbes pour chaque test indiquant le pourcentage de réussites aux différents âges dans toute l'étendue du test. Ce travail fut fait séparément pour les deux sexes et les tests qui favorisaient l'un ou l'autre sexe furent éliminés. La corrélation de chaque test avec le résultat global fut établie pour éliminer les tests les moins valides.

La standardisation juste d'une échelle d'âge dépend du niveau d'âge où sont placés les tests particuliers et aussi du crédit, c'est-à-dire

du nombre de mois d'âge mental, accordé pour réussir le test. D'autres facteurs sont l'intercorrélation des tests et la forme des courbes de pourcentage de réussites. Il est impossible de donner des règles à suivre pour que l'échelle fournisse un quotient intellectuel moyen de cent à chaque niveau. Il a fallu que Terman procède d'une façon empirique en revisant et en revisant de nouveau jusqu'à ce qu'il arrive à un arrangement des tests qui atteignait ce but. Six révisions de la formule L furent nécessaires avant d'arriver à ce résultat. En plus de déplacer les tests, il fallait modifier la façon de les coter pour les rendre plus faciles ou plus difficiles pour qu'ils fassent bien à un niveau déterminé. Ce n'est qu'après avoir fait tout ce travail qu'il présenta les deux formules L et M du New Revised Stanford-Binet.

Il ne suffisait pas à Terman de présenter les deux échelles mais il lui fallait encore prouver leur valeur. Ce sont les caractéristiques techniques qui nous renseigneront sur sa valeur. Terman consacre donc tout un chapitre à l'analyse statistique de ses données. Les deux principales qualités métrologiques d'un test sont la constance et la justesse. Il y a deux mesures de la constance d'un test: le coefficient de constance et l'évaluation de la stabilité des cotes individuelles. Terman présente des coefficients d'équivalence entre les formules L et M qui vont de 0,98 pour les sujets qui ont un quotient intellectuel de 70 ou moins jusqu'à 0,90 pour les sujets qui ont un quotient de 130 et plus. De pareils coefficients témoignent d'un fort degré de précision.

L'évaluation de la stabilité des cotes individuelles, deuxième mesure de la constance d'un test, se fait par le calcul de l'erreur de mesure $\sigma_{(M)}$ ou l'E.P._(M). Pour le New Revised Stanford-Binet cette erreur de mesure est de 1,49 pour les quotients intellectuels de moins de 70; de 2,60 pour les quotients 70 - 89; de 3,04 pour les quotients 90 - 109; de 3,29 pour les quotients 110 - 129; et de 3,54 pour les quotients de 130 et plus². Ceci veut dire que pour les quotients de 90 - 109, par exemple, les variations possibles sont les suivantes: un résultat quelconque a 50 chances sur 100 de ne pas s'éloigner de la vraie mesure de plus de 3 points, ni de moins de 3 points; ce même résultat a 82 chances sur 100 de ne pas s'éloigner de la vraie mesure de plus de 6 points, ni de moins de 6 points; ce résultat a 95 chances sur 100 de ne pas s'éloigner de la vraie mesure de plus de 9 points, ni de moins de 9 points; enfin, il est pratiquement impossible qu'il erre dans un sens ou dans l'autre de plus de 12 points. Ces erreurs de mesure témoignent d'un fort degré de précision.

La deuxième qualité primordiale d'un test est la justesse. Terman emploie deux critères pour juger de la justesse de son échelle³: 1° l'augmentation des pourcentages de réussites d'un âge (ou d'un âge mental) à l'autre, et 2° une pondération basée sur le rapport critique (D/σ_D) entre l'âge moyen (ou l'âge mental moyen) des sujets qui réussissent le

2 Terman, Measuring Intelligence, op. cit., p. 46.

3 Terman, op. cit., p. 9.

test et des sujets qui échouent. Puisque cette pondération est basée sur le nombre total des réussites et des échecs sur l'item en question, il accorde une meilleure base pour juger de la validité relative des items que la série de pourcentages de réussites.

Une publication subséquente de McNemar⁴ fait l'analyse statistique complète des items du New Revised Stanford-Binet. On peut trouver dans ce livre tous les renseignements nécessaires sur la justesse des items. Il y a sans doute certains items qui sont un peu moins valides (ceci arrive dans toute échelle) mais en général il n'y a personne qui mettrait en doute la validité du New Revised Stanford-Binet.

Comme le dit bien Wechsler, c'est l'expérience qui nous montrera si notre test mesure bien ce qu'il prétend mesurer:

How do we know that our tests are good measures of intelligence? The only honest reply we can make is that our experience has shown them to be so. If this seems to be a very tenuous answer we need only remind the reader that it has been practical experience which has given (or denied) final validity to every other intelligence test. Regrettable as it may seem, empirical judgments here as elsewhere, play the role of ultimate arbiter. In any case, all evidence for the validity of a test, whether statistical or otherwise, is inevitably of an indirect sort and, in the end, cumulative rather than decisive⁵.

Ce que Wechsler dit plus haut de la justesse d'un test pourrait faire croire qu'il n'attache aucune importance aux méthodes statistiques.

Alors il ajoute en note au bas de la page:

4 Quinn McNemar, The Revision of the Stanford-Binet Scale, Boston, Houghton-Mifflin, 1942, 185 pages.

5 David Wechsler, The Measurement of Adult Intelligence, Baltimore, The Williams & Wilkins Co., 1944, p. 127.

The reader should not interpret this remark as an attempt on the part of the author to belittle either the importance or value of proper statistical methods. On the contrary, he is of the opinion that without intelligent and thorough statistical evaluation of one's data, little trust can be placed in one's results. The extensive statistical work done on the Bellevue Scales is perhaps the best evidence for that. But he is of the opinion that the task of validating a scale does not end but only begins with the statistical evaluation⁶.

En effet, l'expérience a montré que le New Revised Stanford-Binet est un des meilleurs tests individuels pour les âges de deux à seize ans puisqu'il est le plus connu et le plus utilisé par les psychométriciens.

Cet exposé sur la standardisation du New Revised Stanford-Binet a été fait pour rendre le contraste plus frappant entre le travail de Terman et celui d'Amoss. Après avoir exposé assez sommairement le travail qu'a fait Terman, demandons-nous maintenant ce qu'a fait Amoss ou plutôt ce qu'il n'a pas fait. On trouve absolument rien sur sa standardisation dans son manuel. Comment a-t-il choisi ses items? Quelle étude a-t-il fait des items? Quelle constance, quelle justesse et quel pouvoir discriminatoire peut bien avoir son test? Quelle est l'erreur de mesure du test? Ce sont autant de questions qui restent sans réponse; cependant, ce sont les réponses à ces questions qu'on devrait trouver dans son manuel pour permettre de poser un jugement sur la valeur du test. A supposer qu'on accepterait son test tel quel, les résultats obtenus n'auraient aucune signification puisqu'il est impossible d'en faire une interprétation juste, ne connaissant pas l'erreur de mesure

6 loco cit.

du test,

Il est vrai que le C.I.E. est une imitation du Stanford-Binet puisque la plupart des tests sont tirés presque textuellement de ce dernier. Les modifications qu'Amoss y fait les rendent tantôt plus faciles et tantôt plus difficiles de sorte qu'il est possible que les résultats ressemblent d'assez près aux résultats obtenus sur le Stanford-Binet. C'est ce que fait remarquer le Dr Stothers dans son article intitulé: "The Use of the Canadian Intelligence Test"⁷. Cet article est la seule publication qui donne quelques renseignements sur le C.I.E. et encore sont-ils très sommaires. Le Dr Stothers nous dit trois choses au sujet du test: 1^o qu'Amoss a suivi le plan habituel d'avoir un nombre déterminé de tests à chaque niveau d'âge, 2^o que l'échelle de cotation permet des valeurs d'un, de deux, de trois et de quatre mois pour coter les réponses des sujets, et 3^o que les tests de dessin, d'arithmétique, de vocabulaire et de langage s'étendent à plus de niveaux d'âge que dans d'autres tests d'intelligence. Cependant, il n'explique pas comment le Dr Amoss justifie cette façon de procéder⁸.

Pour prouver la constance du C.I.E., il nous dit qu'Amoss s'est assuré le concours d'un grand nombre d'examineurs dignes de confiance, ceux qui détenaient un Ontario Auxiliary Class Certificate, pour recueillir 1,099 copies du C.I.E. et du Stanford-Binet selon un plan déterminé. Chaque

⁷ C.E. Stothers, The Use of the Canadian Intelligence Examination, dans The School, Elementary Edition, vol. 30, n^o 10, pp. 920 - 923, voir Appendice 2, pp. 177 - 183.

⁸ Loco cit.

examineur devait donner les deux tests à huit élèves de la première à la huitième année inclusivement. On leur faisait suivre un plan déterminé pour choisir les sujets au hasard de façon à inclure des élèves de toutes les origines raciales. Le temps écoulé entre les deux examens ne devait pas dépasser deux jours. Enfin, l'ordre d'administration des deux tests était renversé dans la moitié des cas.

Les coefficients de corrélation tirés de groupes différents sur les deux tests sont respectivement 0,928, 0,924, 0,927, 0,924, 0,925, 0,925, et 0,923⁹. Il conclut de ces résultats que le C.I.E. est un examen constant qui peut servir dans les écoles d'Ontario. Ces coefficients sont très élevés il est vrai, mais le Dr Stothers ne nous dit pas s'ils ont été corrigés par la formule Brown-Spearman et il ne donne pas non plus l'erreur probable de ces coefficients. Ces quelques données statistiques sont insuffisantes pour permettre d'affirmer que le test jouit d'une grande constance.

Voici en quels termes le Dr Stothers résume la discussion de la justesse: "By reason of Section IV of the Regulations for Auxiliary Classes, it became necessary to subject the Canadian Intelligence Examination to a test of validity."¹⁰ C'est assez mince comme renseignements et ça ne prouve pas grand'chose. En réalité, Amoss s'est-il donné la peine de choisir de trois à quatre cents items pour les étudier sous les aspects d'ambiguïté, de réponses possibles, de justesse? A-t-il pris le temps d'analyser les items

9 Loco cit.

10 Loco cit.

particuliers pour chercher à en connaître: 1° la difficulté, c'est-à-dire le pourcentage de réussites aux différents âges, 2° le degré d'accord entre les différents sous-tests d'un même âge, et 3° le point auquel les items distinguent bien entre les élèves qui sont très intelligents et ceux qui le sont moins? Ce travail de sélection exige une habileté technique considérable et beaucoup d'expérience. Il doit être fait avec beaucoup de soin sans quoi le test peut-être facilement gâché à ce stage-ci. Une fois les items sélectionnés, Amoss les a-t-il essayé sur de nouveaux sujets et a-t-il répété tout le procédé pour s'assurer que sa sélection finale était satisfaisante? Il semblerait plutôt qu'Amoss s'est contenté de copier les items du Stanford-Binet et qu'il attribue à ses propres items la même validité que ceux de Terman malgré les modifications qu'il y apporte.

Si, de plus, on se sert de l'argument de validité de Wechsler basé sur l'expérience avec le test, on se rend compte que le test n'a pas beaucoup de valeur puisqu'il a eu très peu de vogue parmi les psychologues. Si le test avait eu une valeur réelle, il y aurait eu des recherches qui auraient été faites. Depuis six ans que le test est sur le marché personne s'est préoccupé d'en étudier la valeur parce qu'il était évident qu'il n'en valait pas la peine. Cependant le Wechsler-Bellevue qui a été publié un an seulement avant le C.I.E. en est à sa troisième édition et il a déjà fait le sujet de nombreuses recherches. Il en a été ainsi pour les révisions de l'échelle métrique faite par des auteurs tels que Terman et Kuhlmann.

Le Dr Stothers présente les résultats de 1,099 élèves choisis avec soin d'une population scolaire de 44,000 élèves d'Ontario. Les données du Tableau I¹¹ et les représentations graphiques de la figure 1¹² de l'Appendice 2 pourraient faire croire à première vue que ces résultats se rapprochent beaucoup de la distribution normale et que par conséquent on peut en conclure que le test d'Amoss est une mesure constante et valide de l'intelligence. Il est à noter que cette courbe a été obtenue en changeant la façon de coter, c'est-à-dire en arrêtant le test à l'année où le sujet n'obtient que des crédits partiels. Le seul motif qui fait agir Amoss est d'ajuster sa courbe pour qu'elle se rapproche davantage de la courbe normale. Ces résultats paraissent assez bien globalement mais rendent-ils justice aux sujets en particulier? Dans le Tableau I le Dr Stothers donne la moyenne, la médiane et le sigma de tout le groupe mais il néglige de faire connaître les moyennes, les médianes et les sigmas de chaque âge. Ce sont autant de renseignements qui permettraient de juger de l'exactitude de la standardisation.

De plus, comment est-il possible d'interpréter les résultats obtenus sur le C.I.E. puisque Amoss ne nous donne pas l'erreur de mesure de son test? C'est une chose qu'il faut absolument connaître sans quoi les résultats obtenus n'ont aucune signification.

Enfin, le Dr Stothers présente les résultats obtenus sur des populations d'écoles primaires et il n'est aucunement question des tests aux deux extrémités de l'échelle. Que valent les résultats obtenus aux âges de

11 Op. cit., p. 182.

12 Op. cit., p. 183.

trois à six, et aussi de douze à adulte supérieur? Puisqu'aucune étude en a été faite il est permis de douter de la valeur d'un bon nombre de ces tests, d'après l'étude que nous en avons faite dans les deux chapitres précédents.

Nous avons donc vu dans ce dernier chapitre tout ce qui manque au C.I.E. pour en faire une oeuvre scientifique de valeur. La présentation du test aussi bien que le matériel laisse à désirer. Amoss a omis de faire connaître les renseignements strictement nécessaires pour permettre de juger la valeur de son test c'est-à-dire les qualités métrologiques du test. De plus, il omet de faire connaître l'erreur de mesure du test sans laquelle toute interprétation des résultats est impossible. Il est donc permis de conclure que l'instrument que présente Amoss pour mesurer l'intelligence est d'une utilité fort douteuse, si on veut rendre justice aux sujets examinés.

CONCLUSION

L'étude que nous avons faite du C.I.E. nous a permis de constater que la plupart des items, au fait cinquante sur les soixante-dix-huit, ont été pris presque textuellement du Stanford-Binet. Amoss ne fait que changer quelques mots. Son principal changement est dans la façon de coter les items; il accorde des crédits partiels à quarante sous-tests. Ces crédits partiels changent la difficulté des items en les rendant tantôt plus faciles et tantôt plus difficiles comme nous l'avons constaté dans l'étude des séries d'items et des items particuliers.

On peut supposer, puisqu'il n'en parle pas, qu'Amoss s'est cru dispensé de faire tout le travail que comporte la standardisation d'un test. Il semblerait bien qu'il part du principe que ses items ayant été en grande partie tirés de la révision Stanford sont nécessairement valides et que ce travail n'a pas besoin d'être fait. Du moment qu'Amoss se permet de modifier les items, ils ne deviennent plus directement comparables à ceux du Stanford-Binet et la standardisation s'impose. En somme, c'est ce qu'a fait Terman dans le New Revised Stanford-Binet. Il prend un grand nombre de sous-tests de son ancienne révision mais il ne se croit pas pour cela dispensé d'une nouvelle standardisation. Au contraire, il l'a fait avec plus de soin que jamais car il comprend bien que c'est la standardisation qui donne la valeur à un test.

Puisqu'Amoss ne donne aucun renseignement sur la standardisation de son échelle et que des renseignements trop insuffisants sont donnés dans

l'article de Dr Stothers¹, il est permis de croire que si cette standardisation a été faite, elle a été faite très sommairement. Une correspondance personnelle avec l'auteur² nous a référé au Dr Stothers qui était censé avoir fait l'analyse statistique du C.I.E. A la demande que nous avons faite à ce dernier³ de nous donner accès aux statistiques du C.I.E. il nous a référé à son article: "The Use of the Canadian Intelligence Test"⁴. Les données qui s'y trouvent sont insuffisantes pour permettre de juger de la valeur du test comme nous en avons conclu dans le dernier chapitre. Il est donc impossible d'obtenir plus de renseignements sur le C.I.E. Ne connaissant pas les qualités métrologiques du test, il devient d'une utilité fort douteuse.

En terminant cette étude nous citons Jackson qui résume dans quelques lignes la conclusion que nous tirons de notre étude et qui s'applique très bien dans le cas du C.I.E.

A carefully constructed test, with a manual of directions and norms, therefore, represents the work of at least three or more persons for at least two years. It will cost, in addition to printing, several thousands of dollars. Not all tests are constructed as carefully as this, but all the better ones are. Ten years of work was spent, and wisely spent, for example, on the last revision of the Stanford-Binet Scales of Intelligence. Since it is a crime to use inferior tests when working with children, test users should examine the literature carefully and select for use only those tests which have been carefully and scientifically constructed.⁵

1 Appendice 2.

2 Correspondance personnelle de l'auteur, lettre du Dr Amoss datée du 16 déc. 1943.

3 Correspondance personnelle de l'auteur, lettre du Dr Stothers datée du 28 août 1945.

4 Stothers, op. cit.

5 R.W.B. Jackson, The Construction of Standardized Tests, dans The School Guidance Worker, vol. 1, n^o 3, livraison de novembre 1945, p. 1-2.

Puisque l'auteur n'a pas fait un travail sérieux de standardisation, et cette tâche lui revient, il est peu probable que quelqu'un entreprenne jamais cette tâche formidable. Nous ne connaissons donc jamais les qualités métrologiques du C.I.E. Il se peut fort bien que le test jouisse d'une certaine constance et d'une certaine justesse, mais n'en connaissant pas le degré ce test devient un instrument de valeur inconnue que le temps fera bien d'ensevelir dans l'oubli.

BIBLIOGRAPHIE

Amoss, Harry, Canadian Intelligence Examination, Toronto, The Ryerson Press, 1940, 69 pages.

Nous y trouvons un avant-propos de l'auteur, une description et des illustrations du matériel de son test, et enfin les directives détaillées des sous-tests. C'est le texte qui fait le sujet de la présente étude.

Amoss, Harry, Ontario School Ability Examination, a performance test prepared more especially for use among children who are deaf, whose native tongue is other than English or who for any reason are lacking in language facility. Toronto, The Ryerson Press, 1936, 54 pages.

Nous y trouvons six genres de tests répétés dans le C.I.E.: 1) le pliage de feuilles, 2) les formes géométriques du Stanford-Revision, 3) le jeu de patience, 4) les poids de Drever et Collins, 5) l'imitation des figures de Gesell et de Stanford, et 6) les dessins de mémoire.

Amoss, Harry, Mathematical Ability Test, Toronto, The Ryerson Press, 2 pages.

Ce test a servi aux séries mathématiques des âges XIV-XVII et XVII-XX.

Gesell, Arnold L., The Mental Growth of the Pre-School Child, New-York, Macmillan, 1925, 447 pages.

Ce livre a servi à Amoss pour quelques tests des premiers âges.

Kuhlmann, F.A., A Handbook of Mental Tests, Baltimore, Warwick & York, 1922, 314 pages.

Les tests de pliage de feuilles sont tirés de ce texte. Il nous sert aussi de point de comparaison.

McNemar, Quinn, The Revision of the Stanford-Binet Scale, An Analysis of the Standardization Data, Boston, Houghton Mifflin, 1942, 187 pages.

Manuel détaillé sur l'analyse statistique des données du New Revised Stanford-Binet.

Stothers, C.E., The Use of the Canadian Intelligence Test, dans The School, Elementary Edition, vol. 30, n° 10, livraison de juin 1942, p. 920-923.

Terman, Lewis M., The Measurement of Intelligence, an explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale, Boston, Houghton Mifflin Co., 1916, 362 pages.

Notre livre de chevet puisqu'Amoss l'a suivi pas à pas.

Terman, Lewis M. et Maud A. Merrill, Measuring Intelligence, a guide to the administration of the New Revised Stanford-Binet tests of intelligence, Boston, Houghton Mifflin Co., 1937, 461 pages.

Ce texte donne plus de renseignements statistiques tandis que le Measurement of Intelligence donne plus de renseignements psychologiques. Comme instrument le nouveau Terman est préférable car il est plus objectif, mais dans les mains de bons psychologues le Measurement of Intelligence a peut-être une plus grande valeur.

Wechsler, David, The Measurement of Adult Intelligence, Baltimore, The Williams & Wilkins Co., 1941, 258 pages.

Ce manuel présente une échelle pour mesurer l'intelligence des adultes. Il sert à faire la comparaison et la discussion de certains tests.

Whipple, Guy Montrose, Manual of Mental and Physical Tests, Part II, Complex Processes, a book of directions compiled with special reference to the experimental study of school children in the laboratory or classroom, Baltimore, Warwick & York, Inc., 1921, 328 pages.

Un chapitre de ce texte renseigne sur les expériences et les recherches faites sur les dessins au miroir.

APPENDICE 1

THE CANADIAN INTELLIGENCE EXAMINATION

FOREWORD

APPENDICE 1

CANADIAN INTELLIGENCE EXAMINATIONS¹

FOREWORD

I

During the past twenty years the Stanford Revision of the Binet-Simon Intelligence Examination has been used successfully by the Auxiliary Classes Branch of the Department of Education, Ontario, in determining the school ability of pupils, who, by reason of physical or mental disability, inferior or superior endowment, seem worthy of special educational consideration. The recent publication by Dr. Terman of a modified and very much elaborated revision, coupled with a possibility that the original edition may be discontinued, gives concern lest an effective, inexpensive and easily administered form of intelligence test should no longer be available to teachers. The present, therefore, appears an opportune time, not only to provide for the continuance of what experience has proved to be a readily applied and reasonably reliable examination, but likewise to iron out certain existing discrepancies and to adjust the content more precisely to the needs of both rural and urban communities.

A tendency in recent years to confuse the functions of the group and the individual test, and to employ the former in situations requiring the use of the latter, constitutes a second reason for the present undertaking. The use of the group test is limited to a study of groups. It

¹ Harry E. Amoss et Charles E. Stogdill, Canadian Intelligence Examination, Toronto, The Ryerson Press, 1940, pp. 5 - 8.

is a valuable means of securing for purposes of comparison general pictures of such aggregates as school, schoolroom, or school-area populations. But results obtained by group testing are not valid for individuals within the group and may be strikingly invalid in case of nervous and overly deliberative candidates. Where placing, grading, promoting, segregating, or studying pupils from the viewpoint of mental hygiene are concerned individual examinations which permit of adjustment in administrative procedure alone should be employed. Hence, it would seem imperative there be available some simple, effective form of individual examination which teachers can learn to apply successfully after a brief course of instruction.

II

In the present work adherence has been given to the general structure of the original Stanford Revision and all modifications have been standardized in relation with that examination. The chief changes are:

(1) Language of instruction used in the presentation of various problems has been simplified.

(2) Attempt has been made more rigorously to define acceptable responses.

(3) Certain tests in which the element of chance seemed unduly involved have been dropped.

(4) With a view to the avoidance of local and literary expressions, the word list in the vocabulary test has been selected from conversational language common to both rural and urban communities rather than from a

dictionary.

(5) To reduce the weighting heretofore given acquired literary skills, more mathematical and constructional problems have been introduced.

(6) Partial answers have been allowed credits for the purpose of obtaining a more finely graded result as well as for the purpose of eliminating over-leniency in the interpretation of responses.

(7) The mean age of mental maturity has been fixed at 15 instead of 16. This change was made in consideration of the fact that practically all retests of children at 16 and 17 who had been previously tested at 8 - 15 showed a falling off in the intelligence quotient. The use of 180 as a maximum denominator largely rectifies this diminution and tends to keep the I. Q. constant.

III

The purpose of the examination is to estimate present intelligence, native or acquired.

To plan truly and avoid fumbling is surely the mark of a superior mind, whether it be engaged in writing an essay or digging a ditch, marshalling troops to victory or organizing a New Year party. An inferior type of mind muddles through, or merely muddles, because: (i) it is unable to take in facts bearing on a situation; (ii) it cannot hold in grasp ascertained data while formulating a plan of action; (iii) it cannot assemble and organize detail into an effective performance picture. A manufacturer neglects to take into account the probability that a political disturbance in Asia may shut off tungsten supplies. A schoolboy forgets to include three days of grace in his calculations while solving

a problem in bank discount. A miner sinking a shaft is unable to figure out the pick-up of a faulted vein.

Intelligence, therefore, may be defined negatively as ability to eliminate trial and error forms of behaviour, and positively as ability to grasp the facts of a situation and use these to formulate plans adequate to occasion. In brief, intelligence is ability to plan.

Both intelligent and non-intelligent people may fail in action owing to lack of special skill or lack of mental integrity. A lad may make an unshapely bird house because he cannot saw straight. A father may bungle a surgical operation on his child because of a trembling hand. A student may abandon a nearly completed project because of jumpy attention. The findings of an intelligence test may indicate but do not afford an estimate of degrees of skill or degrees of mental integration.

Your authors are of the opinion that even within its own field intelligence testing is as yet in the empiric stage of development. Its findings are at best mere approximations, whose numerical indices have no exact mathematical significance, and whose sole value lies in the diagnostic and prognostic uses to which they may be put in school and hospital as determined and confirmed by experience. Intelligence testing will only become an exact science when there shall have been evolved a measuring unit formulated in terms of ability to establish certain mathematically defined relations among certain mathematically defined data, and which yard stick shall be as applicable in estimating the size of a job as in determining the capability of the worker. Nevertheless, after ten years

of labour in the fields of special education and mental health, during which time they have had the privilege not only of examining several thousand children, but likewise of following up the results of advisements based upon these examinations, your authors are also of the opinion that no other appliance has been evolved in the realms of education and psychiatry which compares with the efficiency of the individual intelligence examination as an instrument in practical and scientific investigation.

IV

The Canadian Intelligence Examination is set forth as valid in the Province of Ontario, Canada, and probably in other comparable English-speaking areas for the purpose of estimating the intelligence, and more particularly the school ability, of normal children between ages 3 - 16, superior children between ages 3 - 14, and subnormal persons above age 4.

APPENDICE 2

THE USE OF THE CANADIAN INTELLIGENCE TEST

APPENDICE 2

THE USE OF THE CANADIAN INTELLIGENCE TEST¹

After twenty years of practical experience in intelligence testing and equating achievement with mental age under several elementary school courses of study, Dr. H. E. Amoss, in association with Dr. C. G. Stogdill, prepared the Canadian Intelligence Test, which was published in 1940 by the Ryerson Press. The sub-tests were tried out a part at a time during the administration of the Terman Intelligence Test in urban and rural surveys in Ontario. The test was checked by Dr. Stogdill by administering this examination to 750 pupils.

The Canadian Intelligence Examination is the first individual intelligence test for general use that has been produced in Canada by a Canadian research worker. The Ontario School Ability Examination, published by Dr. Amoss in 1936, is designed for use with subjects who are deaf and have no facility with the English language.

By reason of Section IV of the Regulations for Auxiliary Classes, it became necessary to subject the Canadian Intelligence Examination to a test for validity. It is a compact outfit, consisting of a book of directions and a package of cards. It has the unusual virtue of being inexpensive. The original Terman Intelligence Test had been in use for nearly a quarter of a century and had been revised and improved. Forms L and M of the

¹ C. E. Stothers, The Use of the Canadian Intelligence Test, dans The School, vol. 30, n^o 10, livraison de juin 1942, p. 920 - 925.

Terman-Merrill Examination were too bulky for rural survey work.

The arrangement of the Canadian Intelligence Test follows the usual plan of having a specific number of tests for each age-level. The scoring scale allows values of 1, 2, 3, or 4 months for the candidate's responses. Tests in drawing, arithmetic, vocabulary, and language are extended into more age-levels than in other intelligence tests.

In Ontario schools, the persons who use intelligence tests are the public school inspectors who have conducted surveys of their areas and the teachers of auxiliary classes who have been trained at the summer course in auxiliary education. Approximately 50,000 test reports from inspectors and teachers are on file in the Department of Education and some thousands of re-tests testify to the accuracy of their work. At post-graduate courses in the great universities of the United States, holders of Ontario Auxiliary Class certificates receive credit for a degree and are admitted to educational clinics without the required preliminary courses.

The services of these reliable examiners were enlisted to assemble 1,099 pairs of intelligence tests records according to the following plan:

(a) Each examiner was asked to give the Canadian Intelligence Examination and the Terman Intelligence Test to eight pupils, one each in grades I - VIII inclusive.

(b) A mathematical plan was provided whereby each examiner selected by chance the pupils to be tested.

(c) The examiners were chosen in such a way that pupils of all racial origins would be included.

(d) The time elapsing between the two examinations was not to exceed two days; the order of administration of the two tests was reversed in half the cases.

The test records were received and subjected to the same checking which is required under Section IV of the Regulations for Auxiliary Classes in the cases of pupils recommended for admission to auxiliary classes.

Correlations of the intelligence quotients were worked out at various groupings containing approximately 300, 400, 500, 700, 1,000, and 1,099 pupils, with coefficients respectively of .928, .924, .927, .924, .926, .925, .923.

The comments of the examiners and the checking of each test showed that the Canadian Intelligence Examination was not administered as quickly as the more familiar Terman Intelligence Test. The test records showed that the examination was unduly long because the examiner had to give fourteen to twenty-four different tests in which the pupil received part scores and failures.

The test records were re-scored on the basis of the following variant of the rule for conducting and terminating the examination. "The examination should begin at the year nearest the estimated mental ages given by the teacher and work backward to a year in which the candidate makes a full score and forward to a year in which the candidate makes no score, or a year in which the candidate makes only partial scores; a full year to be construed as any of the sections up to and including year IX - X and either the first or last half of each of the years designated X -XII, XII - XIV, XIV, XIV - XVII, XVII - XX."

Graph I has been prepared to show the relationship of this curve to the theoretical curve.

Table I shows the distribution of the Intelligence Quotients obtained.

Summary. This group of 1,099 pupils is a careful cross-section of approximately 44,000 Ontario pupils. The range of intelligence quotients obtained shows that all types are represented.

The technique of collecting and handling the data is a practical application of the test to our needs, because annually the inspectors of auxiliary classes have had to check upwards of 2,000 intelligence test records as required by the regulations. To be thoroughly scientific the testing should be done by one person with a longer period between examinations. In this case, all the checking has been done by one person.

The results as reported by the examiners show results a little closer than those found in the Rochester Study of 800 pupils tested on the old Terman Test and the Terman-Merrill Form L.

Conclusion. Since the statistics show a consistently close relationship on the two examinations, this leads to the conclusion that the Canadian Intelligence Test is a reliable examination for use in Ontario schools.

The rule for terminating the examination may be varied by ending the test when a full year of partial successes is reached.

The Canadian Intelligence Examination should be valuable as a survey instrument for its definite gradation of learning levels in language, vocabulary, arithmetic, and reading.

Table I

Frequency distribution of 1,099 pupils according to the results of the Stanford Revision of the Binet-Simon Test and the Canadian Intelligence Examination, with and without the revised rule for terminating the examination.

Interval	<u>Stanford Revision</u>	<u>Canadian Intelligence</u>	
	<u>Binet-Simon</u>	<u>Examination</u>	Revised Rule
155-159	1	1	1
150-154	0	0	0
145-149	0	0	0
140-144	3	3	3
135-139	4	5	0
130-135	10	11	12
125-129	26	33	22
120-124	47	48	49
115-119	90	106	90
110-114	97	120	111
105-109	126	146	134
100-104	170	141	156
95- 99	186	170	167
90- 94	125	133	147
85- 89	101	91	96
80- 84	59	46	61
75- 79	29	26	28
70- 74	16	10	12
65- 69	6	7	7
60- 64	3	2	3
55- 59	0	0	0
Totals	<u>1,099</u>	<u>1,099</u>	<u>1,099</u>
Median	100.22	101.78	100.41
Mean	100.86	102.22	100.92
S.D.	13.60	13.35	13.45

APPENDICE 2

183

PUPILS

200
180
160
140
120
100
80
60
40
20
0

65 70 80 90 100 110 120 130 140 150

— actual
- - - theoretical

INTELLIGENCE QUOTIENT

Graph I

Intelligence quotients of 1,000 pupils, obtained by revising the rule for terminating the examination, shown with the theoretical distribution.

APPENDICE 3

FEUILLE DE RAPPORT

DU

CANADIAN INTELLIGENCE EXAMINATION

(PLEASE DUPLICATE ON FORM A.C. 4, ENCLOSED)
CANADIAN INTELLIGENCE EXAMINATION

(Please fill in all blanks and underline characteristic words)

TEACHER'S CONFIDENTIAL REPORT

Municipality _____ School _____ Room No _____
 Pupil's Name _____ Age on first entering a school _____
 Date of Birth _____ Day _____ Mth _____ Yr _____ Present Grade _____
 Nationality _____ Mean Grade Age _____
 Parent or Guardian _____ Retardation _____ yrs _____ mos _____
 Address _____ Grades repeated _____
 Mental Age (teacher's estimate) _____ yrs _____ mos I Q (teacher's estimate) _____
 School Opportunities—good or poor, due to change of school—home conditions _____
 Attendance—regular or irregular, due to truancy—sickness—distance from school _____
 Tendencies—Stable, unstable Restless, quiet Good-natured, ill-tempered Social, isolated _____
 Plays with juniors, age groups, hangs about adults _____

Cleanliness	good, average, poor	Posture	erect,	stooped,	slovenly
Tidiness	good, average, poor	Speech	normal,	defective,	foreign
Politeness	good, average, poor	Alertness	quick,	slow,	dull
Obedience	good, average, poor	Dexterity	handy,	fair,	awkward
Industry	good, average, poor	Reading	Kn, Gr 1, Gr 2, Gr 3, Gr 4, Gr 5, Gr 6, Gr 7, Gr 8		
Self-control	good, average, poor	Writing	Kn, Gr 1, Gr 2, Gr 3, Gr 4, Gr 5, Gr 6, Gr 7, Gr 8		
Kindness	good, average, poor	Arithmetic	Kn, Gr 1, Gr 2, Gr 3, Gr 4, Gr 5, Gr 6, Gr 7, Gr 8		
Truthfulness	good, average, poor	Spelling	Kn, Gr 1, Gr 2, Gr 3, Gr 4, Gr 5, Gr 6, Gr 7, Gr 8		
Honesty	good, average, poor	Drawing	Kn, Gr 1, Gr 2, Gr 3, Gr 4, Gr 5, Gr 6, Gr 7, Gr 8		

REMARKS

Date _____ Grade Teacher _____

NURSE'S CONFIDENTIAL REPORT

Birth—Normal, abnormal _____
 Home Conditions—Uplifting, depressing Comfort, poverty Care, neglect _____
 Brothers and Sisters—older _____ younger _____ school progress _____

Family	}	Health	father	mother
		Mentality	father	mother
		Occupation	father	mother
		Remarks	father	mother

Child	}	Walked at _____	Talked at _____
		Past health _____	
		Present health _____	
		Nutrition _____	
		Personal habits _____	
		Vision _____	Hearing _____
		Speech _____	Crippling _____ Nurse _____

EXAMINER'S CONFIDENTIAL REPORT

Examinee—Nervous, composed Repressed, free Language handicap _____
 Examination—disturbed, uncompleted, abbreviated, satisfactory, unsatisfactory _____
 C A _____ yrs _____ mos M A _____ yrs _____ mos I Q _____
 Recommendation _____
 Date _____ Examiner _____

OTHER EXAMINATION (if any)

C A _____ yrs _____ mos M A _____ yrs _____ mos I Q _____
 Date _____ Examiner _____

REMARKS

YEAR II-III

- | | |
|--|-----|
| 1. Copying Circle..... | Sc. |
| 2. Copying Cross..... | |
| 3. Folding—1..... 2..... | |
| 4. Naming—1..... 2..... 3..... 4..... 5..... | |
| 5. Picture—1..... 2..... 3..... | |
| 6. Digits—312..... 231..... 213..... | |
| 6A. Syllables—1..... 2..... 3..... | |

YEAR IV-V

- | | |
|--|--|
| 1. Copying Triangle..... | |
| 2. Colours—1..... 2..... 3..... 4..... | |
| 3. Folding—1..... 2..... | |
| 4. Meanings—1..... 2..... 3..... | |
| 4..... 5..... 6..... | |
| 5. Completion—1..... 2..... 3..... | |
| 6. Commissions..... | |

YEAR VI-VII

- | | |
|---|--|
| 1. Copying Diamond..... | |
| 2. Counting Fingers—1..... 2..... 3..... | |
| 3. Pictures—1..... 2..... 3..... | |
| 4. Digits—31759..... 42835..... 96817..... | |
| 5. Differences—1..... 2..... | |
| 3..... 4..... | |
| 6. Digits B—283..... 427..... 958..... | |
| 6A. Days—1, 2, 3, 4, 5, 6, 7. Tu..... Sa..... Th..... | |

YEAR VIII-IX

- | | |
|---|--|
| 1. Change—1..... 2..... 3..... | |
| 2. Definitions—1..... 2..... 3..... 4..... | |
| 3. Digits B—6258..... 4937..... 8692..... | |
| 4. Sentences—1..... 2..... 3..... | |
| 5. Rhymes—1..... 2..... 3..... | |
| 6. Months—J, F, M, A, M, J, J, A, S, O, N, D..... | |
| Ap..... Ju..... Sep..... | |
| 6A. Weights—1..... 2..... 3..... | |

Picture Descriptions:

- | |
|---------|
| 1. |
| 2. |
| 3. |
| 4. |

YEAR III-IV

- | |
|---|
| 1. Copying Square..... |
| 2. Counting..... |
| 3. Folding—1..... 2..... |
| 4. Forms..... |
| 5. Comprehension—1..... 2..... 3..... |
| 6. Digits—4135..... 2514..... 4253..... |
| 6A. Syllables—1..... 2..... 3..... |

YEAR V-VI

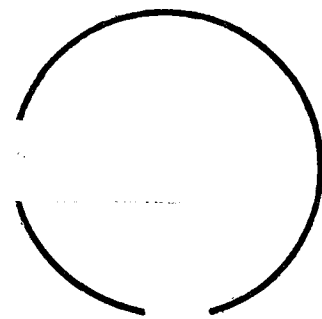
- | |
|---|
| 1. Copying Pentagon..... |
| 2. R.L.—1..... 2..... 3..... 4..... 5..... 6..... |
| 3. Coins—5..... 1..... 25..... 10..... |
| 4. Counting..... |
| 5. Omissions—1..... 2..... 3..... 4..... |
| 6. Syllables { 1..... |
| 2..... |
| 3..... |

YEAR VII-VIII

- | |
|--------------------------------|
| 1. Copying Star..... |
| 2. Counting B..... |
| 3. Likenesses—1..... 2..... |
| 3..... 4..... |
| 4. Stamps—1..... 2..... 3..... |
| 5. Vocabulary—10..... 8..... |
| 6. Drawing—1..... 2..... |
| 6A. Dictation..... |

YEAR IX-X

- | |
|--|
| 1. Vocabulary—20..... 17..... |
| 2. Ratio— $\frac{10 \times 6}{3}$ $\frac{20 \times 8}{4}$ $\frac{5 \times 6}{2}$ |
| 3. Drawing—1..... 2..... |
| 4. Absurdities—1..... 2..... |
| 3..... 4..... 5..... |
| 5. Words..... |
| 30..... 24..... |
| 6. Digits—374859..... 521746..... 837592..... |



YEARS X-XII			Sc.
Vocabulary—30.....28.....25.....			
Ratio— $\frac{12 \times 5}{3}$ $\frac{24 \times 7}{4}$ $\frac{56 \times 9}{7}$			
Digits B—31879,69482,52961.....			
Dollar and Field.....			
Definitions—1.....2.....			
3.....4.....5.....			
Sentences	1.....		
	2.....		
	3.....		
	4.....		
Pictures—1.....2.....			
3.....4.....			
Likeness—1.....2.....			
3.....4.....5.....			
YEARS XIV-XVII			
Vocabulary—50.....47.....44.....			
Drawing—1.....2.....3.....			
Differences—1.....2.....			
3.....4.....5.....6.....			
Digits B—471952,583294,752638.....			
Arithmetic—1.....2.....3.....4.....			
Series.....			

YEARS XII-XIV			Sc.
1. Vocabulary—40.....38.....36.....34.....			
2. Induction—1.....2.....3.....			
4.....5.....Rule.....			
3. Digits—2183469.....9628475.....			
4. Arithmetic—1.....2.....			
3.....4.....			
5. Clock—6.22.....8.08.....			
4.57.....2.46.....			
6. Difference—1.....2.....			
3.....4.....5.....			

YEARS XVIII-XX			
1. Series.....			
2. Arithmetic—1.....2.....3.....			
3. Differences—1.....2.....			
3.....4.....			
5.....6.....			
4. Digits B—4162593,3826475,9452837.....			
5. Reading—1.....2.....3.....			
6. Mirror—1.....2.....3.....			

1	4	3	6	5	8	7	10
1	2	4	7	11	16	22	29
4	5	7	11	19	35	67	131
11/12	5/6	3/4	2/3	7/12	1/2	5/12	1/3
2.15	2.5	2.85	3.2	3.55	3.9	4.25	4.6
2	5	9	19	37	55	109	219
1	3	8	19	42	89	184	375
1/6	2/3	1/3	5/6	1/2	1	2/3	11/6
3	7	12	27	50	105	204	415

_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

VOCABULARY

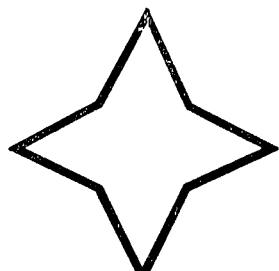
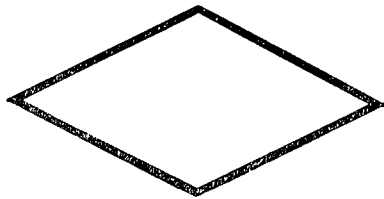
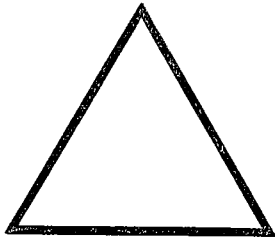
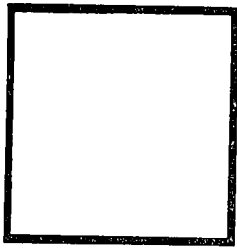
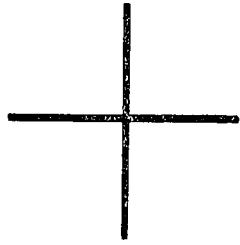
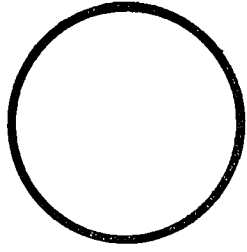
ple
ddle
velope
aw
pper
er
ith
e
lash

dangerous
absent
difficult
peculiar
return
hesitate
narrow
false
outward
southern
consider

value
lengthen
familiar
profit
recent
loiter
skill
gradual
insure
expand
control

brunette
former
succeed
desolate
diminish
bemoan
priceless
inferior
conscientious
vertical
tolerate

maximum
convex
optimistic
corpulent
superfluous
typical
frustrate
intermittent
tentative
deprivation
avarice



APPENDICE 4

AN ABSTRACT OF.

Une étude critique et comparative

du

Canadian Intelligence Examination

method of giving partial credits in forty sub-tests whereby the difficulty of the tests is either increased or decreased, thus making it impossible to compare them directly with those of the Stanford Revision.

Fundamentally, the changes in the wording, in the presentation, and in the scoring of the sub-tests do not matter providing the author proves the validity of these items as set forth in his scale. However, Dr. Amoss has omitted this necessary information in his manual. The only information available on which to judge the value of the test is a four-page article entitled: "The Use of the Canadian Intelligence Test" by Dr. C. Stothers in the School, Elementary Edition of June 1942. The statistical data set forth in this article seems insufficient to prove the reliability and the validity of the Canadian Intelligence Examination.

Until proof of the contrary can be found, we feel that the Canadian Intelligence Examination lacks adequate standardization. Its total validity and its item validity, as much as its reliability, are unknown. Hence its value as an adequate intelligence measuring instrument is very doubtful.

Appended to the body of the Etude are the following:

1. The Canadian Intelligence Examination - Foreword (Amoss introduction to the Canadian Intelligence Examination.)
2. The Use of the Canadian Intelligence Test (a four-page report of the Canadian Intelligence Examination by Dr. C.E. Stothers.)
3. A sample of the Canadian Intelligence Examination Record Blank.
4. An abstract of Une étude critique et comparative du Canadian Intelligence Examination.

F I N I S

APPENDICE 4

AN ABSTRACT OF

Une étude critique et comparative du Canadian Intelligence Examination¹.

The Canadian Intelligence Examination published by Dr. Harry E. Amoss and Dr. Charles G. Stogdill in 1940 and set forth as a valid general intelligence test for the Province of Ontario has as yet received little experimental control. Since this test is being used in the school of Ontario it was thought appropriate to make a detailed, critical and comparative study of the test in all its parts.

A discussion of his presentation, of his purpose, and of his concept of intelligence introduces the thesis. A detailed study of all the sub-tests made from the standpoint of their sources, the procedure and the directions to be followed, the modifications made, the scoring suggested was accompanied by a comparison with similar tests of Binet, Terman and Kuhlmann. Finally, a judgment was carried on the value of the sub-tests proposed by Amoss. A last chapter was assigned to the critical study of the test as a whole: its general presentation, its material and mainly its reliability and its validity.

Our study revealed that fifty of the sub-tests were taken almost literally from the Stanford Revision with only slight changes made in the wording, the more important changes being in the scoring. Amoss adopts the

¹ Eva Labrosse, Une étude critique et comparative du Canadian Intelligence Examination, thèse présentée à la faculté des arts de l'Université d'Ottawa par l'intermédiaire de l'Institut de Psychologie, Ottawa, 1946, IX - 188 pages.