CHAPITRE V

RECHERCHE D'UNE MÉTRIQUE COMMUNE AUX QUATRE CAHIERS DE TEST : ÉTALONNAGE EN QUOTIENT D'INTELLIGENCE

par Pierre Benedetto

L'évaluation de l'intelligence d'un enfant à l'aide de l'Échelle Collective de Niveau Intellectuel est déterminée par la réussite ou l'échec à chacune des diverses questions de l'épreuve. Il paraît naturel de résumer l'ensemble de ces résultats par une « note » obtenue en faisant la somme des bonnes réponses données par l'élève.

Si toutes les questions constituant l'épreuve sont convenablement choisies, il paraît également normal de décréter « plus intelligent » le sujet qui réussit 21 questions que celui qui ne fournit que 20 bonnes réponses.

Cependant, il n'est pas possible d'étendre à l'ensemble des symboles numériques ainsi obtenus les propriétés arithmétiques des nombres. En effet, lorsque le biométricien mesure la taille de deux enfants et qu'il obtient 0,60 m pour l'un et 1,20 m pour l'autre, nul ne contestera l'assertion suivante : le deuxième enfant a une taille deux fois plus grande que le premier. Lorsque le psychologue évalue l'intelligence de deux individus et que le premier a répondu correctement à 30 questions d'un test d'intelligence et que le second a réussi 60 questions du même instrument peut-on dire que le deuxième sujet est deux fois plus intelligent que le premier? Certainement non.

Pour s'en convaincre, il suffit d'imaginer que l'on ajoute à l'épreuve 100 questions très faciles réussies par les deux sujets; la note du premier devient 130, celle du second 160. Le rapport entre la note du premier et celle du second sujet est devenu très différent.

Or, le rôle assigné aux mesures psychologiques nécessite l'utilisation d'opérations arithmétiques. Par exemple, pour l'Échelle Collective de Niveau Intellectuel, le dépouillement de l'enquête fera intervenir le calcul de moyennes, de variances, d'indices de corrélation. Toutes ces opérations supposent implicitement que l'on traite une grandeur mesurable. Une grandeur est dite mesurable d'après Campbell (Reuchlin (M.) *Mesures quantitatives en psychologie*, p. 16) à condition :

1° Qu'une relation d'ordre puisse être établie entre les choses à mesurer.

2° Partant de ce que l'addition constitue le fondement des propriétés arithmétiques des nombres, qu'une opération expérimentale isomorphe à l'opération d'addition puisse être réalisée sur cette grandeur.

L'utilisation de l'Échelle Collective de Niveau Intellectuel est donc subordonnée à la détermination d'une métrique qui remplisse les conditions ci-dessus. D'autre part la constitution de cette échelle en 4 cahiers séparés pose des problèmes particuliers : la métrique doit être la même pour tous les cahiers.

Nous allons examiner successivement comment il est possible de déterminer une métrique au sein d'un même cahier puis comment on passe à une métrique commune pour les 4 cahiers successifs.

I. — RECHERCHE D'UNE MÉTRIQUE AU SEIN D'UN MÊME CAHIER

1. Ordination des sujets. Un cahier est constitué par plusieurs sous-tests. Chaque sous-test est composé d'une série de questions élémentaires formellement comparables par leur contenu. Ces questions ont été classées suivant un ordre de difficulté croissante. La difficulté d'une question a été estimée au moyen de l'enquête pilote. Le test a été passé alors par un échantillon de sujets comparable à celui de l'enquête définitive et la difficulté est définie par la « proportion de sujets ne parvenant pas à donner une réponse correcte après un temps de travail suffisamment long pour que l'on puisse supposer que son prolongement ne modifierait pas le résultat ».

On fait alors l'hypothèse suivante : il existe un caractère sous-jacent qui permet de résoudre les questions et, tel qu'à chaque niveau de réussite correspond un degré de ce caractère. Il devient alors facile de classer les individus suivant leur intelligence à partir des réponses positives à l'Échelle Collective de Niveau Intellectuel. L'individu classé premier est celui qui a résolu la question la plus difficile et ainsi de suite.

L'ordination des sujets ainsi réalisée, nous devons trouver dans l'expérience une opération qui satisfasse à l'axiomatique de l'addition.

2. De l'ordination des sujets à la comparaison des intervalles. Dans les mesures physiques, on compare l'intervalle qui sépare deux objets à un étalon déterminé extérieurement à l'expérience et on « compte » combien de fois l'étalon est compris dans l'intervalle. L'intervalle est ainsi comparé à une *somme* d'intervalles égaux à l'étalon.

En psychologie, il n'en est pas ainsi; l'étalon n'est pas défini hors de l'expérience. Comment alors comparer les intervalles séparant les individus? Comme il y a correspondance univoque entre les individus et le degré de difficulté que ceux-ci sont parvenus à résoudre, le problème revient à comparer les degrés de difficultés entre eux.

Mais les intervalles de difficultés ne sont pas directement comparables. Rien, en effet, ne nous permet de comparer la différence de difficultés entre deux questions résolues l'une par 50 % des sujets et l'autre par 40 % et la différence de difficultés entre deux questions résolues l'une par 20 % des sujets et l'autre par 10 %.

Pour trouver une solution à ce problème, le psychologue est conduit à faire une hypothèse qui lui permettra de définir l'égalité entre deux intervalles

Dans le domaine des mesures physiques, celui de la biométrie par exemple, il est d'usage courant de dresser à partir des données relevées sur un groupe d'individus des courbes appelées histogrammes. L'utilisation de ces graphiques peut se faire de deux manières : ou bien étant donnée la valeur de la mesure pour un individu, on détermine son rang; ou bien, étant donné le rang d'un individu, on recherche sa note. La connaissance des rangs de deux individus nous permet de déterminer leurs positions respectives dans l'histogramme et la distance qui sépare ces positions nous renseigne directement sur la différence entre leurs mesures.

Dans le domaine des mesures psychologiques, l'inexistence d'étalon antérieur à l'expérience ne permet pas de déterminer des intervalles entre les mesures et, de ce fait, il n'est pas possible de déterminer l'histogramme des mesures. L'hypothèse du psychologue porte sur la forme de cet histogramme. Cette hypothèse formulée, nous nous retrouvons dans le deuxième cas d'utilisation des histogrammes relatifs aux mesures physiques : connaissant les rangs respectifs de deux individus, il est aisé de les situer dans l'histogramme et la distance qui sépare leurs positions dans l'histogramme traduit directement la distance entre leurs mesures respectives. Les intervalles compris entre les mesures deviennent alors comparables. En déformant une distribution expérimentale et en la rendant conforme à un modèle (sans modifier la relation d'ordre entre les sujets) on peut diviser l'axe des abscisses en une série d'intervalles égaux. Ce procédé porte le nom de *normalisation* lorsque le modèle est la courbe normale de Laplace Gauss. Il sera décrit en détail plus loin.

3. De la comparaison des intervalles à l'addition.

Si on désire situer un sujet quelconque sur l'axe des abscisses, on peut effectuer à l'aide du test plusieurs estimations successives. Il est certain que ces estimations ne seront pas toutes les mêmes. La meilleure de toutes ces estimations est obtenue en prenant la moyenne arithmétique de cet ensemble, ce qui nous conduit à effectuer une *addition*.

Une autre façon de procéder consiste à effectuer plusieurs estimations, non pas en répétant les applications d'un même test, ce qui entraîne des variations systématiques dues à l'effet d'apprentissage, mais en effectuant plusieurs mesures à l'aide de tests censés mesurer la même dimension psychologique ou bien à l'aide de tests mesurant divers aspects de cette dimension. Or, l'Échelle Collective de Niveau Intellectuel correspond à ce dernier cas; l'intelligence générale d'un individu est estimée par *sommation* des scores dans les différents sous-tests. Nous avons donc trouvé, dans l'expérience, une opération qui satisfait à l'axiomatique de l'addition.

Le lecteur pourra ici présenter l'objection suivante : dans le cas de l'Échelle Collective de Niveau Intellectuel aucune normalisation n'est effectuée à partir de la distribution des scores bruts par sous-tests. On normalise les distributions des notes verbales et des notes non verbales, notes obtenues en faisant la somme des bonnes réponses à chacune des parties correspondantes (voir chapitre II) sans apparemment se préoccuper de la forme des distribution par sous-tests.

En fait, les sous-tests ont été construits de manière à maintenir à peu près constante la différence de difficulté entre les items consécutifs.

Voici, brièvement exposé le procédé [1] que nous avons utilisé pour réaliser cette équidistance.

On utilise un diagramme pour lequel l'axe des abscisses représente les degrés de difficulté; on lit sur l'axe des ordonnées le nombre d'individus qui parviennent au degré de difficulté correspondant à un point de l'axe des abscisses. On fait alors l'hypothèse suivante : la forme de cet histogramme est celle de la courbe normale de Laplace-Gauss.

Prenons un exemple. Dans sa version définitive, on décide que tel sous-test contiendra 15 items. Ces 15 items sont séparés par 14 intervalles que nous désirons égaux.

On considère 15 valeurs centrées et équidistantes de la variable normale centrée réduite z; on prendra les suivantes par exemple (d'autres pourraient être choisies) :

$$-\frac{21}{10}\quad -\frac{18}{10}\quad -\frac{15}{10}\quad -\frac{12}{10}\quad -\frac{9}{10}\quad -\frac{6}{10}\quad -\frac{3}{10}\quad 0\quad \frac{3}{10}\ \frac{6}{10}\ \frac{9}{10}\ \frac{12}{10}\ \frac{15}{10}\ \frac{18}{10}\ \frac{21}{10}$$

La table de la loi normale réduite donne alors la proportion d'éléments de l'ensemble infini normal qui n'atteignent pas la valeur de z correspondante.

---

(1) On trouvera un développement plus détaillé de ce principe dans les ouvrages suivants : ANASTASI A. « Psychological Testing ». New York. *Mac Millan*, 1954, chapitre 7;
THORNDIKE R. L. «Personnel Selection : Tests and Measurement Techniques». New York. *Wiley*, 1949, chapitre 8.

A chaque point ainsi défini, correspond un degré de difficulté et ces degrés de difficulté sont équidistants.

Lors de la construction de la version définitive, nous nous sommes efforcé de ne retenir que les items dont le degré de difficulté (évalué à l'aide de la version expérimentale lors de la pré-enquête) était voisin de celui des points théoriques définis ci-dessus. Ce procédé rend licite l'*addition* du nombre de bonnes réponses à chacun des sous-tests afin d'obtenir les notes verbales et non verbales.

Nous voyons donc à partir de quelles hypothèses et selon quels principes, on parvient à transformer les notes brutes d'un test en une série de mesures sur lesquelles il est licite d'effectuer des opérations arithmétiques.

II. — L'INFLUENCE DE L'ÂGE : NOTION D'ÂGE MENTAL

1. Cas d'un cahier unique. Dans tout ce qui précède, nous avons admis que le caractère mesuré par l'épreuve est seul responsable de la réussite ou de l'échec à une question. Dans le cas d'une échelle de développement intellectuel, la réussite (ou l'échec) à une question est influencée par l'âge de l'individu. Il est des questions qui sont réussies à un certain âge, d'autres à un âge plus avancé. C'est cette considération qui a entraîné le découpage de l'Échelle Collective de Niveau Intellectuel en 4 cahiers consécutifs. A chaque cahier correspond un niveau scolaire donc une certaine distribution des âges. Mais tous les enfants d'un même âge n'obtiennent pas le même résultat au test et réciproquement, parmi les individus qui ont la même note, tous n'ont pas le même âge. Au cours du dépouillement et de l'exploitation des résultats de l'enquête de 1944, en vue de permettre la comparaison de niveaux d'intelligence observés par des sujets d'âges différents, les scores bruts au test Mosaïque de Gille ont été présentés en *âge mental.* Un sujet, par exemple, a 9 ans d'âge mental s'il obtient le résultat moyen observé dans la population des enfants normaux de 9 ans.

Dans le cas où le test est composé en cahiers successifs, on ne peut plus procéder ainsi. Quel sens aurait, en effet, la moyenne des notes obtenues par des enfants qui n'ont pas passé le *même* cahier?

2. Cas où le test est structuré en plusieurs cahiers. Il semble bien évident que la situation de l'expérimentateur dans ce cas est très différente de celle qu'il connaît lorsque le test est composé d'un cahier unique. Or, cette différence n'est qu'illusoire. Replaçons-nous, pour démontrer cette proposition dans l'hypothèse d'un cahier unique. Des élèves de 6 ans et des élèves de 14 ans ont passé ce même cahier de test. Les élèves de 6 ans ont été classés sur une plage de questions; les élèves de 14 ans sur une autre plage de questions. La distance entre les deux plages est appréciée par la somme des poids affectés à chaque élément d'un ensemble de questions. Cet ensemble de questions se décompose en un ensemble de plages successives non disjointes deux à deux classant les élèves de 7 à 13 ans. Or, nous avons vu que l'étendue de ces plages peut être modifiée par le constructeur, par exemple lorsqu'il veut améliorer la sensibilité de son épreuve. Donc, la distance entre deux plages est purement arbitraire. De ce fait, la courbe décrivant la variation de la note à un test en fonction de l'âge est également arbitraire et Zazzo[1] a dénoncé l'absurdité des thèses émises par certains auteurs qui prétendaient obtenir, à partir des notes brutes d'un test, une courbe de croissance mentale comparable à celle obtenue pour le poids et la taille.

Si la distance entre les notes moyennes des enfants d'âge a, d'une part, et les enfants d'âge b, d'autre part, est dépendante de l'instrument de mesure, il en est, à fortiori, de même pour la distance entre la note x d'un enfant A d'âge a et celle y d'un enfant B d'âge b. Il faut, cependant, comparer A et B. Pour ce faire, nous allons comparer les distances de A et B aux enfants « moyens ayant respectivement le même âge que A et B. C'est d'ailleurs ainsi qu'on procède lorsque dans le langage courant (celui des mères de famille) on parle d'enfants « en avance », d'enfants « en retard ». La référence au groupe d'enfants du même âge est alors implicite. Si Pierre et Paul sont tous deux « en avance », on dira que Pierre qui a 14 mois, est moins en avance que Paul qui a 8 mois si la distance qui sépare Pierre de l'enfant moyen de 14 mois est moins grande que la distance qui sépare Paul de l'enfant moyen de 8 mois.

Pour que les distances ainsi définies soient comparables, il faut qu'elles soit évaluées à l'aide d'échelles ayant la même unité. Quant à l'origine, sa position étant arbitraire, elle peut, de ce fait, être choisie une fois pour toute et maintenue fixe quel que soit l'âge.

La fixation arbitraire de la référence n'est pas en contradiction avec le phénomène de croissance mentale. Un exemple tiré de la vie scolaire permettra d'illustrer ce principe. Lorsqu'un chef d'établissement décide qu'il faut pour être admis dans la classe supérieure, une moyenne égale ou supérieure à 9,5 (ou bien qu'il est nécessaire, pour avoir le prix d'excellence, d'avoir une note moyenne supérieure à 15), il adopte pour toutes les classes une échelle commune (moyenne 10, unité = 1 point). Il considère que deux élèves A et B ayant la même moyenne sont comparables quel que soit leur niveau scolaire. Ainsi si A et B ont 9,5 comme note moyenne ils sont comparables : ce sont les enfants « ayant acquis le bagage juste nécessaire pour entrer dans la classe supérieure ». Si A est en classe de sixième et B en première, nul ne songerait à attribuer à A et à B le même savoir.

Dans sa révision de l'Échelle Métrique d'intelligence de Binet et Simon, le choix de l'unité de mesure a posé de grands problèmes à Terman [2].

---

[1] Zazzo R. « Le devenir de l'intelligence ». Paris, *P.U.F.*, 156 pages.
[2] Terman L. M. « The measurement of intelligence ». Londres. *Harrap*, 1919, 461 pages.

Nous avons dit que chaque groupe d'enfants homogène quant à l'âge, est classé sur une plage de questions. Terman a cherché à construire ses plages de questions telles que les distributions des notes obtenues pour tous les groupes homogènes quant à l'âge soient identiques c'est-à-dire de même forme et définies par les mêmes paramètres. Or la forme de la distribution des notes obtenues à un test provient uniquement d'un choix délibéré du constructeur. Il est donc licite de procéder d'une manière inverse à celle adoptée par Terman. Au lieu de rechercher, à priori, une suite de distributions ayant toutes une même forme et mêmes paramètres, on détermine expérimentalement les distributions brutes et, à posteriori, on les transforme afin qu'elles aient la même forme et les mêmes paramètres.

Le procédé que nous venons de décrire a été adopté par Wechsler [1]. C'est ainsi qu'il détermine le *quotient intellectuel*.

III. — LE QUOTIENT INTELLECTUEL

Le Q.I. (Quotient intellectuel) est souvent défini comme le rapport entre l'âge mental et l'âge civil tous deux représentés en mois. Une telle définition masque le caractère statistique du Q.I. qui cependant est réel.

On considère que l'âge mental d'un enfant est normal, c'est-à-dire égal à son âge civil lorsque cet enfant obtient au test une note égale à la moyenne des notes obtenues par l'ensemble des sujets de son âge civil.

De la même manière, on dit qu'un enfant à un âge mental de x mois lorsqu'il obtient une note égale à la moyenne des notes obtenues par l'ensemble des sujets ayant un âge civil de x mois.

Donc l'âge mental d'un enfant n'est autre chose que sa note dans le test, note qu'on a choisi d'exprimer en unités appelés mois.

Lorsqu'on calcule un Q.I., ce score est rapporté à la note exprimée en mois que l'enfant aurait obtenue s'il avait été identique à l'enfant *moyen* du même âge civil.

On a donc :

$$\text{Q. I.} = \frac{\text{A. M.} \times 100}{\text{Age réel}} = \frac{\text{Note obtenue par le sujet} \times 100}{\text{Note moyenne obtenue par les enfants du même âge}}$$

On voit maintenant le parallélisme entre le Q.I. et l'échelle de mesure dont nous avons parlé précédemment. Le rapport $\frac{\text{AM}}{\text{AR}}$ n'est qu'un indicateur de la distance qui sépare un individu de l'individu moyen du groupe d'individus du même âge.

Restent seulement à définir le point de référence de l'échelle et son unité.

---

(1) WECHSLER D. « La mesure de l'intelligence de l'adulte ». Paris, *P.U.F.*, 1956, 292 pages.

TABLEAU V.1. — NORMALISATION DE LA DISTRIBUTION DES NOTES TOTALES
CAHIER I. COURS PRÉPARATOIRE

Notes brutes (x)	Fréquence des notes	Fréquence cumulée		Variable gaussienne		
		Observée F (x)	Ajustée	Centrée réduite $m = 0$; $\sigma = 1$	Note normalisée $m = 100$ $\sigma = 15$	Note arrondie
0	0,2	0,23		— 2,8782	56,8	57
1	0,3	0,53		— 2,5758	61,4	61
2	0,6	1,16		— 2,2571	66,1	66
3	0,6	1,77		— 2,0969	68,5	69
4	0,9	2,64		— 1,9431	70,9	71
5	0,9	3,58		— 1,7991	73,0	73
6	1,3	4,93		— 1,6546	75,2	75
7	1,5	6,39		— 1,5220	77,2	77
8	1,4	7,76		— 1,4187	78,7	79
9	2,0	9,78		— 1,2930	80,6	81
10	2,3	12,12		— 1,1700	82,4	82
11	2,3	14,41		— 1,0625	84,1	84
12	2,5	16,92		— 0,9581	85,6	86
13	2,5	19,46		— 0,8596	87,1	87
14	3,2	22,69		— 0,7488	88,8	89
15	3,1	25,79		— 0,6495	90,3	90
16	3,3	29,08		— 0,5505	91,7	92
17	3,0	32,10		— 0,4649	93,0	93
18	4,0	36,13		— 0,3558	94;7	95
19	4,3	40,44		— 0,2430	96,4	96
20	4,3	44,79		— 0,1307	98,0	98
21	4,6	49,43		— 0,0150	99,8	100
22	5,2	54,63		+ 0,1156	101,7	102
23	4,9	59,50		+ 0,2404	103,6	104
24	5,2	64,70		+ 0,3772	105,7	106
25	4,7	69,42		+ 0,5072	107,6	108
26	4,8	74,22		+ 0,6495	109,7	110
27	4,5	78,70		+ 0,7961	11,9	112
28	4,3	82,96		+ 0,9542	114,3	114
29	3,5	86,50		+ 1,1031	116,5	117
30	3,4	89,86		+ 1,2759	119,1	119
31	2,8	92,63		+ 1,4466	121,7	122
32	2,5	95,10		+ 1,6546	124,8	125
33	1,7	96,80		+ 1,8522	127,8	128
34	1,4	98,20		+ 2,0969	131,5	131
35	0,9	99,07		+ 2,3656	135,5	135
36	0,5	99,60		+ 2,6521	139,8	140
37	0,3	99,89		+ 3,0902	146,4	146
38	0,1	100,00	99,95	+ 3,30	149,5	149
39	0,0	100,00	99,98	+ 3,55	153,3	153
40	0,0	100,00	99,99	+ 3,80	157,0	157
	100,0					

La référence de l'échelle de mesure est égale à 100, valeur du Q.I. lorsque l'individu se confond avec l'individu moyen de son groupe d'âge.

Avant de définir l'unité de mesure, il convient de faire une hypothèse sur la forme de la distribution. Partant de l'expérience quotidienne qui montre que peu d'individus sont soit nettement supérieurs, soit nettement inférieurs à la moyenne, on a *arbitrairement* choisi comme fonction de densité de la loi du Q.I. celle de la loi de Laplace-Gauss.

Quant à l'indice de dispersion, il est conventionnellement choisi de telle manière que les 50 % des sujets médians de chaque âge soient compris entre les valeurs 90 et 110 de Q.I. Il est d'usage courant, en effet, de fixer la frontière entre les individus d'intelligence moyenne et ceux au-dessous (ou au-dessus) de cette moyenne, d'une part à une fois l'erreur probable et, d'autre part, à un Q.I. égal à 90 (ou 110). On démontre alors aisément que l'écart-type de la variable Q.I. vaut environ 15 points.

Afin de déterminer le Q.I., l'expérimentateur doit utiliser la distribution des scores x pour chaque groupe d'enfants du même âge.

Si la distribution des scores x n'est pas normale, on la rend normale selon le procédé appelé *normalisation*.

IV. — NORMALISATION D'UNE DISTRIBUTION

Voici matériellement comment on réalise une telle transformation à partir des résultats expérimentaux obtenus au cahier I appliqué au cours préparatoire. La distribution expérimentale n'est pas normale. La deuxième colonne du tableau V-1 indique la fréquence des notes; on désire normaliser cette distribution expérimentale. La fonction de répartition de la loi normale réduite est donnée par les tables usuelles correspondant à une variable z de moyenne zéro et d'écart-type unité.

A chaque fréquence cumulée observée dans la population on fait correspondre une valeur z lue dans la table de la loi normale.

Lorsque la distribution des scores x est normale (ou rendue normale), pour chaque groupe d'enfants du même âge, on passe de la variable x (m_x, σ_x) à la variable Q.I. ($M_{QI} = 100$, $\sigma_{QI} = 15$) en utilisant la transformation :

$$\text{Q. I.} = \frac{x - m_x}{\sigma_x} \times 15 + 100$$

V. — CAS PARTICULIER D'UNE ÉCHELLE COMPOSÉE DE CAHIERS SUCCESSIFS

Mais, dans le cas particulier d'une échelle présentée en cahiers successifs, la transformation indiquée ci-dessus n'est pas immédiatement possible, tous les enfants du même âge n'ayant pas passé le même cahier.

Il faut donc effectuer une correction qui permette de comparer des enfants du même âge fréquentant des classes différentes. Il est aussi nécessaire de tenir compte des différences d'âge pour des enfants d'une même classe.

Voici un procédé permettant cette correction.

Pour l'ensemble de l'échantillon de l'enquête, on connaît la distribution des notes brutes obtenues dans une classe par les enfants ayant passé le cahier correspondant à cette classe. On en déduit une distribution de la variable x ayant pour toutes les classes, la même moyenne (100) et le même écart-type ($\sigma_x = 15$).

Pour un sous-échantillon expérimental, les enfants d'une classe passent non plus le cahier prévu pour le niveau scolaire mais celui du niveau scolaire immédiatement supérieur ou immédiatement inférieur.

Prenons un exemple. Le cahier II est *normalement* passé par les élèves du cours élémentaire. Pour un sous-échantillon expérimental, ce cahier II est également appliqué à des élèves du cours moyen. Ce dispositif expérimental est responsable de la complexité du plan de sondage.

On transforme les scores bruts des élèves du cours élémentaire, de manière à obtenir une distribution des notes x de moyenne 100 et d'écart-type 15. Si l'on exprime les scores des élèves du cours moyen à l'aide de la même métrique, la distribution sera asymétrique, la moyenne de cette distribution sera également supérieure à 100 ,120 par exemple, l'écart-type aura également varié.

Toutes choses étant égales par ailleurs, la différence entre les scores moyens des élèves du cours élémentaire et celui des élèves du cours moyen est imputable à la différence des âges moyens entre les groupes d'enfants constitués par les deux cours.

Admettons, par exemple, que l'âge moyen des enfants du cours élémentaire soit de 8 ans et que l'âge moyen des élèves du cours moyen soit de 10 ans.

A une différence d'âge de 2 ans correspond une différence de 20 points.

— Jean est élève du cours élémentaire. Il passe le cahier II. Il a 8 ans. Sa note brute obtenue lors de l'application du cahier II transformée en variable x fournit directement son Q. I.

— Paul est élève du cours élémentaire. Il passe le cahier II. Il a 10 ans. Sa note exprimée en variable x est 110. Or, cette note est exprimée dans une métrique correspondante à un âge moyen de 8 ans. Paul a 2 ans de plus. Étant donné qu'à une différence de 2 ans correspond une différence de 20 points de score, on obtient son Q. I. en retranchant 20 points à sa note en x c'est-à-dire : 90.

— Louis est élève du cours élémentaire. Il passe le cahier II. Il a 9 ans. Sa note exprimée en variable x est 115. Si à une différence de 2 ans d'âge correspond une différence de 20 points de Q. I., on peut admettre qu'à une différence d'un an correspond une différence de 10 points de Q. I. Le Q. I. de Louis est donc de 105 points.

On voit maintenant comment on tient compte des variations de l'âge au sein d'une même classe.

En fait, le système de correction est plus compliqué. Nous avons déjà signalé que l'écart-type de la distribution obtenue à partir de l'échantillon expérimental est, en général, différent de 15. La méthode proposée ci-dessus est fondée sur le fait qu'à une variation d'âge correspond une translation globale de la distribution des notes. Ce principe suppose donc l'invariance de l'écart-type. Or il n'en est pas ainsi. Pour tenir compte de cette réalité, nous avons partitionné pour un même cahier la distribution des notes de l'échantillon principal (ensemble des élèves qui ont passé le cahier correspondant normalement à leur âge), et la distribution des notes de l'échantillon expérimental. Dans ces deux partitions, la première partie correspond aux notes inférieures à $(m - 1\ \sigma)$; la seconde, les notes comprises entre $(m - 1\ \sigma)$ et $(m + 1\ \sigma)$; la troisième partie comprend les notes supérieures à $(m + 1\ \sigma)$.

Pour chaque partie, nous avons calculé un point moyen. On considère que la distance entre les points moyens de deux parties correspondantes de l'échantillon principal et de l'échantillon expérimental est imputable à la différence d'âge.

Il est apparu, du fait de certaines dissymétries dans les distributions des âges, que les notes obtenues après les traitements proposés ci-dessus ne se répartissaient pas tout à fait normalement et que les paramètres des distributions n'étaient pas exactement égaux à 100 et à 15.

Une dernière normalisation est donc nécessaire pour obtenir une variable à laquelle on puisse attribuer réellement le qualificatif de Q. I. (1).

Les transformations que nous venons de décrire peuvent être appliquées aux notes partielles obtenues par la totalisation des bonnes réponses aux cahiers II, III et IV. On détermine ainsi un Q. I. verbal et un Q. I. non verbal.

Ces deux notes partielles sont mesurées à l'aide d'un même nombre de questions dans un cahier donné. Mais suivant les cahiers, les notes obtenues ne sont pas identiques dans leurs distributions, leurs moyennes, leurs dispersions.

Il suit que, dans le total de ces deux notes, elles ne jouent pas un rôle égal, l'une pesant plus que l'autre et selon les cahiers, c'est tantôt la note verbale, tantôt la note non verbale qui prédomine.

(1) Au cours de cette dernière normalisation, nous avons fait correspondre les notes observées les plus basses à un Q.I. de 70. Nous avons ainsi tenu compte et réinséré dans la population d'un âge donné un pourcentage approximatif de 2 % d'enfants non scolarisés ou dont le quotient d'intelligence est inférieur à 70. Dans le chapitre I, page 13 on a donné une autre estimation des non scolarisés dont la proportion serait inférieure à 1 %. Nous nous refusons néanmoins d'apprécier à l'aide d'une échelle collective un quotient d'intelligence inférieur à 70.

Pour pallier cet inconvénient, la note globale n'est pas obtenue en faisant la somme des notes brutes verbale et non verbale, mais en faisant la somme des notes obtenues après la normalisation de ces notes brutes. La note globale est donc obtenue quel que soit le cahier en effectuant la somme de deux variables normales de moyenne 100 et d'écart-type 15.

Il convient de souligner, en terminant, le caractère *approximatif* des mesures effectuées à l'aide de cette Échelle. Cette approximation est due au fait que l'examen est collectif. On ne saurait, en effet prétendre obtenir le même degré de précision par un examen individuel et un examen collectif. L'examen individuel est aussi une approximation, mais toutes choses égales par ailleurs, elle est meilleure. On peut cependant penser que, pour l'examen collectif, l'erreur sur la mesure est faible lorsque l'individu examiné n'est pas très éloigné de la normale. C'est le cas de tous les enfants examinés lors de l'enquête de l'I.N.E.D. puisque cette enquête ne porte que sur les enfants scolarisés.

Il conviendra de n'utiliser qu'avec une extrême prudence toute mesure effectuée à l'aide de cette échelle sur les enfants non scolarisés. Dans ce cas l'examen individuel est indispensable.