

## Les deux métriques concurrentes pour créer des scores unifiés.

Pour la construction des scores unifiés, à partir des mêmes résultats calibrés par cahier, sont apparues deux "métriques" concurrentes (métrique INOP, ou métrique INED).

*Tableau 12. Variables de résultats au test : les scores unifiés créés en 1966*

	Optique a-métr.	Optique compré- hensive	Optique psycho- technique
<b>+++++++ SCORES UNIFIÉS (variables J***)</b>			
= Scores unifiés de la <b>métrique INED (JA**)</b>			
- Score interclassant tous les élèves avec test ....	///	JAI*	JAI*
-> Aux âges d'obligation , quotient intellectuel QI	///	JAQI*	JAQI*
-----			
= Scores unifiés de la <b>métrique INOP (JEP**)</b>			
disponibles <b>seulement en champ principal étendu CEP</b>			
- score n°1 dit corrigé pour l'effet d'âge .....	///	JEP1*	JEP1*
-> Aux âges d'obligation , score n°2 déduit de JEP1	///	JEP2*	JEP2*
-----			
Rappel <b>Résultats spécialisés par cahier</b> (Variables K***) voir détails	///	///	K***

### 1. Usage du plan de recherche dans la métrique INED (création du score interclassant).

La métrique INED crée d'abord un score interclassant les cahiers. Les performances au sein d'un cahier donné sont connues par les résultats du calibrage par cahier :

- au cahier II par le score équipondéré dans ce cahier (K2G<sub>z</sub>)
- au cahier III par le score équipondéré (K3G<sub>z</sub>)
- au cahier IV par le score équipondéré (K4G<sub>z</sub>)
- au cahier I par la note brute globale (K1B) ou par le score gaussien global (K1B) ; les options étaient possibles ; dans le cas du cahier I, la métrique INED donne les mêmes résultats, puisque les rangs des élèves selon la note brute sont intégralement conservés par le score gaussien global.

La démarche veut obtenir un score interclassant les performances de tous les élèves, quel que soit le cahier qu'ils ont passé. Les principes de la méthode sont du démographe Louis Henry.

La méthode utilise les duos d'échantillons parallèles disponibles dans une même colonne du plan de recherche. Huit duos sont disponibles dans le tableau 4.

Aux différences près qui découleront des tirages au sort, les deux échantillons d'un duo sont équivalents pour décrire les élèves en France dans un certain niveau scolaire ; la seule différence systématique est l'emploi de deux cahiers de test différents. En CM2, par exemple, les deux échantillons sont équivalents, aux écarts aléatoires près, pour décrire la diversité des élèves de CM2 en France. Admettons le postulat de l'équivalence de contenu psychologique du cahier III et du cahier IV : c'était l'objectif du créateur des cahiers. Avec ce postulat psychologique, les deux échantillons du CM2 expriment les mêmes aptitudes ; seules diffèrent les échelles formelles de mesure.

La réalisation finalement retenue utilise trois duos, dans l'ordre suivant : duo en CM2, duo en CE2, puis duo en CE1 (voir tableau 13).

Tableau 13. Après collecte : usage du plan d'enquête dans la métrique INED pour la création du score interclassant les performances par cahier.

	CP	CE1	CE2	CM1	CM2	CO1	CO2	4e	PRF
						FE1 6e	FE2 5e		
Rappel: Elèves <b>sans test</b> présent :	S	S	S	S	S	S S	S S	S	S
<b>Elèves avec test présent :</b>									
Elèves avec <b>CAHIER 1 présent</b>	tPP	tex	--	--	--	--	--	--	--
Elèves avec <b>CAHIER 2 présent</b>	tEXL	tPP	tPP	tex	--	--	--	--	tPQ
Elèves avec <b>CAHIER 3 présent</b>	--	--	tEX	tPP	tPP	tex	tex	tee	--
Elèves avec <b>CAHIER 4 présent</b>	--	--	--	--	tEX	tPP	tPP	tPP	tPP
Elèves avec <b>CAHIER 4 présent</b>	--	--	--	--	tEX	tPP	tPP	tPP	tPP
Elèves avec <b>CAHIER 3 présent</b>	--	--	tEX	tPP	tPP	tex	tex	tee	--
Elèves avec <b>CAHIER 2 présent</b>	tEXL	tPP	tPP	tex	--	--	--	--	tPQ
Elèves avec <b>CAHIER 1 présent</b>	tPP	tex	--	--	--	--	--	--	--

Le duo d'échantillons en CM2 donne en parallèle deux distributions différentes, mais qui expriment en réalité le même classement des élèves de CM2 en France. Une correspondance bijective peut être établie, centile par centile : on établit ainsi une table de correspondance bijective entre les notes obtenues au cahier III et les notes obtenues au cahier IV. Cette table de correspondance, établie en CM2, permet d'interclasser entre eux, non seulement les élèves du mélange des deux échantillons du CM2, mais tout élève avec cahier III avec tout élève avec cahier IV. Ainsi le niveau scolaire CM2 a servi de premier relais.

Choisissons le CE2 comme second niveau relais (CE2) ; les deux échantillons disponibles dans ce niveau permettront l'interclassement des élèves avec cahier II et des élèves avec cahier III (déjà interclassés avec ceux avec cahier IV). Enfin, en utilisant le CE1 comme troisième relais, le duo d'échantillons dans ce niveau permettra d'interclasser tous les élèves avec cahier I, avec tous les élèves avec cahier II, qui sont déjà interclassés avec tous les autres élèves.

On obtient finalement le score interclassant de la métrique INED, défini pour tous les élèves ayant un test présent, dans toutes les cases du plan d'enquête. Ce score interclassant permet de comparer tous les élèves comme avec un unique cahier (à la manière de l'enquête de 1944), en oubliant totalement quel cahier réel ils ont passé. Les résultats du score interclassant sont évidemment sous l'effet des âges inégaux des élèves.

Cette lecture du plan de recherche par la métrique INED est décrite comme une "méthode de transformation des notes par niveau scolaire" (Tome 2, p.13) ; ce qui est un résumé abusif.

=Problème d'interprétation

Signalons le problème théorique de l'interprétation à donner à un score interclassant tous les élèves. On peut affirmer que le classement des individus, et surtout la distance apparente entre eux, dépend du choix des questions élémentaires (items) choisies par le

créateur du test. Un test n'a aucune objectivité intrinsèque ; ce que souligne Pierre Benedetto (opus INED n°54 p.97-98). Entre l'emploi de quatre cahiers (1965) et un cahier unique (1944), "la différence est illusoire. [...] "la courbe décrivant la variation de la note à un test en fonction de l'âge est également arbitraire, et Zazzo<sup>1</sup> 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". Est-ce une condamnation de tous les tests d'intelligence ?

A quoi on peut répondre qu'un même problème existe en toute description des faits sociaux. Le raisonnement puriste conduirait à s'interdire tout classement socio-professionnel, puisqu'il n'existe aucune objectivité intrinsèque dans la meilleure des nomenclatures ; un code disponible est une délimitation approximative, non totalement arbitraire si on admet que toutes les situations sociales ne sont pas identiques. Concluons ainsi : la croissance de la courbe des performances avec l'âge, arbitraire dans sa forme détaillée, n'est pas totalement arbitraire, mais est une approximation du réel, s'il existe un accroissement des performances avec la montée en âge des enfants. Dans la métrique INED, la création du score interclassant n'était donc pas jugée condamnable.

Les particularités de la métrique INOP ne sont pas détaillées ici. Le projet original de P. Benedetto (opus INED n°54 p.102) suit une méthode nouvelle dont il espérait, en 1965, qu'elle produirait directement un QI, en refusant la création préalable d'un score interclassant. La mise en application a donné la variable que nous nommons "premier score de la métrique INOP", et que P. Benedetto nommait "corrigé de l'effet de l'âge". Un deuxième score en a été tiré, pour donner aux résultats du premier score la forme canonique d'un quotient intellectuel (à chaque âge, distribution gaussienne de moyenne 100 et d'écart type 15).

## 2. Exploitation en mode agrégé : création du quotient intellectuel (métrique INED)

Rappel : un quotient intellectuel est un score qui respecte la forme canonique suivante : les enfants de même âge en France ont une diversité de performance à un test ; cette diversité à âge égal est exprimée par un score respectant la forme d'une distribution en cloche (Laplace-Gauss), de moyenne 100 et d'écart type 15.

Une condition préalable a été strictement respectée : le QI ne peut être obtenu que si l'on dispose au préalable d'une distribution de **tous les enfants de même âge selon leurs performances**.

Les travaux donnant le score interclassant ont été produits par quatre exploitations, séparément par cahier. Pour arriver au quotient intellectuel, tous ces élèves, muni du score interclassant, sont exploités en mode agrégé, c'est à dire sans distinguer les cahiers passés, ni les niveaux scolaires fréquentés.

Les élèves sont ventilés par classe trimestrielle d'âge le jour du test. Chaque classe d'âge montre une distribution particulière selon le score interclassant ; en particulier le score croît en moyenne avec l'âge (Tome 2, p.35). Aux âges d'obligation (6 ans à moins de 14), chaque classe trimestrielle dans l'échantillon décrit tous les élèves scolarisables en France, à peu près. Dans chacune de ces classes trimestrielles avant 14 ans, on construit le quotient intellectuel en imposant la forme canonique (forme gaussienne, moyenne 100, écart type 15)<sup>2</sup>.

---

<sup>1</sup> Zazzo, René, *Le devenir de l'intelligence*, Paris, P.U.F., 156 p.

6..Avec minimum à 65. Les élèves non scolarisables sont supposés être écartés des écoles pour déficience intellectuelle majeure ; tous les non scolarisés (environ 1%) ont donc un QI théorique

*-a) Usage du quotient intellectuel de la métrique INED.*

Quand on met en relation le score interclassant avec d'autres variables, il faut tenir compte des âges des élèves dans les divers groupes comparés. La solution lourde est alors de fractionner l'effectif de chaque groupe en classes d'âge assez homogènes. Telle a été la procédure dans les analyses de l'enquête de 1944.

Pour exploiter l'enquête de 1965, on a prévu d'écartier l'effet de l'âge sur les performances (score interclassant de la métrique INED) en créant un deuxième score unifié, de style "quotient intellectuel". Sa fonction est statistique : donner un classement des performances à âge égal. Ce QI a été utilisé de façon exclusive dans les publications descriptives (tomes 2 et 3). L'avantage pratique est considérable : en qualifiant chaque élève par son QI, et non par le score interclassant, il est inutile de se préoccuper de la diversité des âges des élèves dans les divers groupes comparés. Deux élèves d'âge différents et qui ont des performances brutes différentes peuvent alors être jugés équivalents, une fois que l'on a écarté l'effet attendu des différences d'âge.

En exploitant l'enquête de 1944, la création d'un score de style QI n'avait pas été faite; la possibilité théorique existait, et elle est donnée en exemple en 1959 par Maurice Reuchlin, adoptant la méthode de Wechsler<sup>3</sup>. Dans la décennie 1950, le calcul d'un QI aurait été difficile, à une époque où, sans ordinateur, les exploitations ont été faites en trieuse mécanique. Une autre raison est l'orientation psycho-sociale des analyses faites après 1944 : René Gille voulait non seulement écartier l'effet de l'âge des élèves, mais aussi l'effet de leur appartenance sociale. Deux élèves de même âge mais de milieux sociaux différents, et qui ont des performances brutes différentes, peuvent être jugés équivalents si l'on écarte l'effet attendu de la différence des milieux d'origine. L'enquête de 1965, en utilisant le QI, a montré directement les différences entre catégories socio-professionnelles à âge égal. Mais sans créer un score qui serait "à âge égal et à milieu social égal".

*-b) Le quotient intellectuel, défini seulement de 6 à moins de 14 ans.*

L'usage du score QI entraîne une sélection fondamentale. Alors que le score interclassant est connu pour tout élève avec test exploitable, le calcul strict du QI n'a été possible que pour les élèves en âge d'obligation scolaire le jour du test.

Cette limitation découle d'une définition stricte du quotient intellectuel adoptée dans la métrique INED<sup>4</sup>. Un QI doit montrer le classement de l'individu en référence à la distribution des performances de tous les enfants de son âge en France. Dans le domaine scolaire de

---

inférieur à 65, selon la loi de Laplace-Gauss (de moyenne 100 et d'écart type 15) ; pour les élèves scolarisés (et donc l'échantillon de l'enquête), on a fixé le QI minimum à la valeur 65.

<sup>3</sup> Reuchlin, Maurice, "La définition du 'quotient d'intelligence' ", *Biotypologie*, XX n°1 (mars 1959) p.13-24. -- L'auteur, comme directeur de l'INOP, est co-signataire de l'avant-propos publié en tête de chacun des trois tomes publiés (opus INED n°54, n°64 et n°83). L'auteur du test ECNI, membre de l'INOP, cite la méthode de D. Wechsler (opus n°54 p.99-101), adoptée par M. Reuchlin, mais il ne l'applique pas.

<sup>4</sup> Cette définition stricte impose une condition démographique qui n'est pas formulée par M. Reuchlin dans l'article cité.

l'enquête en 1965, certains élèves sont hors de la période légale d'obligation scolaire : ils ont atteint ou dépassé leur 14e anniversaire et sont restés scolarisés. Aux mêmes âges, d'autres enfants sont déjà sortis de scolarité, et ne sont donc pas représentés dans les distributions du score interclassant. Dès lors, on ne peut établir un score de type QI pour tout élève qui a dépassé 14 ans.

Dès lors, sur l'effectif des 115 006 élèves avec test (et avec un score interclassant dans la métrique INED), une bonne fraction (15 918) est resté sans score de type QI. La raison essentielle est qu'ils étaient hors des âges de scolarité obligatoire à la date du test. Les résultats ventilant le score QI ont donc porté, dans le fichier Fic1, sur 99.088 élèves (effectif Tq en tableau 14). C'est, en gros, le sous-domaine des élèves, avec test, qui avaient de 6 ans à moins de 14 ans le jour du test.

Tableau 14. Effectifs exploités en 1966 dans le traitement du fichier spécialisé Fic1 (nombres retrouvés en fichier fossile 1970: les élèves avec bagage F1 )

Fic1	--->	122.074	<b>élèves avec bagage F1 (étaient présents en fichier Fic1)</b>
(S)=		- 7.068	sans test exploitable (test fantôme, ou test réel non exploitable).
T.....	=	115.006	<b>élèves avec test présent et exploité</b>
			Ils ont un score unifié interclassant en métrique INED
			Ils ont des résultats propres au cahier passé, exploités avec poids mono-cahier
		12.799	élèves <b>avec cahier 1</b> (résultats K1*) avec poids mono-cahier
		25.522	élèves <b>avec cahier 2</b> (résultats K2*) avec poids mono-cahier
		29.574	élèves <b>avec cahier 3</b> (résultats K3*) avec poids mono-cahier
		45.111	élèves <b>avec cahier 4</b> (résultats K4*) avec poids mono-cahier
T_/ Th	=	-15.918	Hors des âges d'obligation : sans score unifié de style QI
\ Tq	=	99.088	<b>élèves de 6 à moins de 14 ans avec un score QI en métrique INED</b>

*-c) Sélection implicite découlant de l'emploi du quotient intellectuel.*

Sur 100 élèves avec score interclassant, 86% entrent dans les conditions d'âge pour avoir un QI calculé. Cette proportion globale masque une grande hétérogénéité, et aussi un grave problème. En utilisant le quotient intellectuel au lieu du score interclassant, l'exclusion des élèves qui ont déjà atteint 14 ans a pour effet de tronquer très inégalement les effectifs de chaque niveau scolaire. La troncature est négligeable dans les premiers niveaux (CP CE1 CE2 CM1 CM2) ou en classe de perfectionnement. Mais dans la description des niveaux scolaires les plus haut (Co1 Co2 4e), la vue de la diversité complète des élèves fréquentant un niveau scolaire est faussée : en utilisant le QI, on se limite aux élèves de moins de 14 ans dans le niveau scolaire.

Si l'on veut la vue complète de la diversité des élèves fréquentant les classes de quatrième, on ne doit pas se contenter d'observer les élèves de moins de 14 ans. Ce serait la sélection des élèves les plus jeunes, ou "en avance". L'usage du QI produit effectivement cette sélection. L'emploi du QI disponible rend impossible de procéder aux comparaisons usuelles, par exemple entre 4e et CM2 : retenant les élèves avec QI, qui ont tous moins de 14 ans, on ne peut plus faire la comparaison de tous les élèves de 4e à tous ceux de CM2. Un piège existe, qui a été insuffisamment marqué dans les publications (toutefois, voir opus INED n°83, p.2, dans l'étude du cycle d'observation).

### 3. Les deux scores unifiés de la métrique INOP

A l'opposé du score interclassant (métrique INED), défini pour tout élève avec test, la

métrique INOP a fourni deux scores unifiés qui ont deux caractères :

1- les scores de la métrique INOP ne sont pas définis pour tous les élèves avec test, mais seulement pour un certain sous-ensemble, formant le champ principal étendu (présenté en tableau 5-A) ;

2- la métrique INOP s'interdit de fournir un score rendant visible un interclassement des performances, comme on le verrait avec un unique cahier fictif.

L'exposé des procédures de construction n'est pas fait ici.

Dans l'idée initiale de P. Benedetto, ce score "corrigé pour l'âge" serait le quotient intellectuel cherché . Nous nommons ce résultat "score n°1".

Sans mettre en doute la validité de la méthode, l'observation des résultats de ce premier score a montré qu'ils ne répondaient pas aux critères canoniques formels d'un QI : dans chaque classe trimestrielle d'âge, le score obtenu n'avait pas une moyenne constante de 100, un écart type de 15, et la distribution ne respectait pas la forme en cloche. Un deuxième score a été créé, imposant aux résultats du premier score le respect des normes canoniques (distribution gaussienne, moyenne 100, écart type 15 ). C'est ce second score qui, avec les aménagements manuels de P.Benedetto, a fondé les tables pragmatiques diffusées pour accompagner l'usage commercial du test ECNI ( publication en Tome 1).