



ESSAI DES ITEMS D'UN TEST PILOTE DE CULTURE GENERALE A KINSHASA (RDC)

Pascal GBAKA NDAYA¹, Florentin AZIA DIMBU², Fidèle NZOKANA TE NZABEMA³, Basile KALEMA KISUNGU⁴

Université Pédagogique Nationale, R.D. Congo

Résumé

La République Démocratique du Congo fait face à un défi majeur lié à la pauvreté en matière de construction de tests psychologiques. Pour remédier à cette situation, nous avons entrepris d'enrichir les testothèques en créant un test de connaissances qui intègre les spécificités culturelles et contextuelles du pays. Dans cette première version, notre attention s'est portée sur l'analyse des items, et les résultats attestent que la majorité d'items est idoine, ce qui confirme notre hypothèse de travail.

Mots clés : Construction de tests, analyse des items, indice de discrimination, indice de difficulté, items idoines.

Abstract

The Democratic Republic of the Congo faces a major challenge related to the lack of resources for the development of psychological tests. To address this situation, we have undertaken to enrich test libraries by creating a knowledge test that incorporates the country's cultural and contextual specificities. In this first version, our focus was on the analysis of items, and the results show that the majority of items are appropriate, confirming our working hypothesis.

Keywords: Test construction, item analysis, discrimination index, difficulty index, appropriate items.

Digital Object Identifier (DOI): <https://doi.org/10.5281/zenodo.18346476>

¹ Candidat au Doctorant en Psychologie Sociale et des Organisations et Assistant-Chercheur en Psychométrie à l'Université Pédagogique Nationale (UPN) de Kinshasa (RDC).

² Professeur des Universités, expert en Psychométrie et en Psychopathosociale à l'Université Pédagogique Nationale (UPN).

³ Chef de travaux à l'Université Pédagogique Nationale et Chercheur en Psychométrie.

⁴ Actuellement Candidat à la thèse, également Chef de travaux à l'Université Pédagogique Nationale, spécialisé en Psychologie Sociale.

1 Introduction

Le test est une épreuve standardisée, définie dans ses conditions d'application et son mode de notation, permettant d'évaluer des comportements ou des processus psychiques d'individus en les comparant à ceux d'un groupe de référence, afin de mesurer des différences individuelles de manière quantitative ou typologique.

Pour qu'un test puisse être utilisé avec confiance, il doit répondre à plusieurs critères essentiels: il doit être fidèle, c'est-à-dire aboutir à la même conclusion lors de plusieurs administrations; valide, c'est-à-dire mesurer effectivement ce qu'il prétend évaluer, et sensible, c'est-à-dire capable de différencier les individus forts de ceux qui sont faibles. En d'autres termes, un test psychologique doit respecter les caractéristiques métrologiques fondamentales. (Rathus 1991; Chartier, P., & Loarer, E., 2008; Azia Dimbu, 2024).

En République Démocratique du Congo(RDC), l'absence d'instruments de mesure adaptés aux réalités locales est principalement due à un manque de formation des spécialistes. Bien que certains experts utilisent des tests occidentaux, cela conduit souvent à des résultats biaisés. Comme le soulignent Mabila Seda et Banzikila Sile (1986), cités par Nzokana Te Nzabema (2011), un Congolais compétent peut perdre une opportunité simplement à cause d'un instrument de mesure basé sur des normes étrangères.

Pour remédier à ce problème, il est conseillé d'adapter les tests étrangers à la réalité locale après une phase d'expérimentation, à condition que l'utilisateur dispose du manuel d'instruction de ces tests. Toutefois, l'approche idéale serait de concevoir des tests locaux ou, en leur absence, de développer de nouveaux tests adaptés au contexte local.

Face à la rareté des instruments psychométriques adaptés aux réalités du Congo, nous travaillons actuellement à l'élaboration d'un test pilote de connaissances visant à évaluer la culture générale des étudiants congolais. Ce test sera soumis à un essai pour vérifier la pertinence et l'adéquation des items sélectionnés. Cette étape nous amène à nous poser les questions suivantes :

- les items du test pilote sont-ils idoines ?
- combien d'entre eux sont inadéquats ?
- quel est le moment le plus propice pour administrer ce test afin d'obtenir des résultats significatifs ?
- cette version du test nécessite-t-elle des ajustements avant sa mise en œuvre définitive ?

2 Approche méthodologique

2.1 Population

La population cible est constituée de l'ensemble des étudiants de l'Université Pédagogique Nationale pour l'année académique 2023-2024. Il est important de noter que cette population est considérée comme infinie en raison de l'absence de données fiables sur le nombre exact d'étudiants. Cette situation est principalement due aux retards observés dans le paiement des frais académiques par certains étudiants, ainsi que des perturbations et décalages éventuels dans le calendrier académique. (Gbaka Ndaya, 2024 : 18).

2.2 Echantillon et ses caractéristiques

Selon D. Laveault et J. Gregoire (2014 : 13), « la mise à l'essai consiste à soumettre l'ensemble des items à un échantillon de la population. Cet échantillon n'a pas besoin d'être représentatif ni d'une taille particulièrement importante. sa dimension dépend essentiellement de l'hétérogénéité de la population ciblée par le test et de la taille de la population de référence ». Cela étant, nous avons sélectionné un échantillon occasionnel de 100 étudiants de première année de LMD de l'Université Pédagogique Nationale, composé de 56% des étudiants et de 44% des étudiantes, afin de procéder à un essai de ce test pilote de connaissances. La sélection d'un échantillon de cette taille est justifiée par la nécessité d'évaluer les items des tests pilotes, ce qui permettra une analyse approfondie et une sélection appropriée des items idoines, comme l'affirme Azia Dimbu (2024: 163).

Tableau 1. Présentation de l'échantillon selon la variable promotion

Promotion	Fréquences
L1LMD Psychologie	35
L1LMD OSP/Jour	8
L1LMD OSP/Soir	10
L1LMD GAISF/Jour	9
L1LMD GAISF/Soir	38
Total	100

Source: Gbaka Ndaya P., mémoire de licence (2024:19)

Le tableau 1 présente la répartition de l'échantillon selon la variable promotion. On constate que 38% des étudiants sont issus de la promotion de L1 LMD GAISF/Soir, suivis par ceux de L1 LMD Psychologie avec 35%. Les étudiants de L1 LMD OSP/Soir représentent 10%, tandis que ceux de L1 LMD GAISF/Jour et L1 LMD OSP/Jour constituent respectivement 9% et 8%.

Tableau 2. Présentation de l'échantillon selon la variable tranche d'âges

Tranche d'âges	Fréquences
Moins de 18ans	9
18ans et plus	91
Total	100

Source: Gbaka Ndaya P., mémoire de licence (2024:19)

L'analyse du tableau 2 révèle que la majorité des étudiants de première année, à savoir 91 %, ont 18 ans ou plus, tandis que les moins de 18 ans, regroupés avec les adolescents, ne représentent que 9 % de l'échantillon. Cette répartition indique que l'échantillon est principalement composé de jeunes adultes, ce qui correspond au profil d'âge habituel chez les étudiants en enseignement supérieur. Elle témoigne également d'une certaine homogénéité sur les plans du développement cognitif et psychosocial.

La prédominance des étudiants âgés de 18 ans et plus (91 %) suggère que l'échantillon est constitué majoritairement de jeunes adultes ayant atteint le stade du développement cognitif formel, tel que décrit par Piaget (1972). Selon cette théorie, à partir de l'adolescence tardive, les individus entrent dans le stade des opérations formelles, caractérisé par la capacité de raisonnement abstrait, hypothético-déductif et logique. Ces compétences facilitent la compréhension des consignes de recherche, l'analyse des situations et la formulation de réponses cohérentes et réfléchies.

La faible proportion d'étudiants de moins de 18 ans, qui constitue 9 % de l'échantillon, peut s'expliquer par des parcours scolaires accélérés ou des entrées précoces dans l'enseignement supérieur. Cependant, malgré leur engagement dans le même stade cognitif que leurs pairs plus âgés (Piaget, 1972), leur développement psychosocial reste en cours de consolidation, ce qui peut influencer la stabilité de leurs opinions et la profondeur de leurs réponses (Piaget & Inhelder, 1966). La forte représentation des jeunes adultes dans l'échantillon contribue à assurer une homogénéité accrue sur le plan cognitif et développemental, renforçant ainsi la fiabilité et la validité interne des résultats. En définitive, cette configuration de l'échantillon permet d'interpréter les résultats principalement à partir des expériences, représentations et capacités réflexives propres aux étudiants ayant atteint un niveau avancé de maturité cognitive, conformément aux postulats du développement intellectuel de Piaget.

2.3 Collecte de données

Un test de connaissance à choix multiple, rédigé en français, a été administré aux étudiants. Conçu pour évaluer la culture générale, il s'agit d'une épreuve écrite (papier-crayon), adaptée au contexte éducatif de la RDC, où le français est la langue officielle.

Le test pilote, élaboré en 2024, comporte initialement 60 items avec 5 propositions chacune, dont une seule correcte. Ce surplus permet de sélectionner les 40 ou 50 meilleurs items après évaluation par des experts, conformément aux recommandations de Laveault et Gregoire (2014).

Le mode d'administration est collectif ou individuel, à temps libre, offrant ainsi à chaque participant la possibilité de répondre à son rythme. Cela, facilite l'analyse des réponses et permet de calculer la durée de la passation du test. Chaque question exige que l'examiné encercle une lettre correspondant à la réponse choisie. La cotation attribue un point par réponse correcte et zéro point aux items non répondus ou incorrects.

La population ciblée comprend des adultes scolarisés détenteurs d'un diplôme d'Etat. Ce type de test, classé parmi les évaluations d'efficiency mentale selon Azia Dimbu (2019), est également reconnu pour sa validité dans le contexte professionnel, selon Lievens et Patterson (2011).

2.4 Analyse des données

Le test pilote de culture générale d'inspiration congolaise a été analysé et traité à l'aide de trois techniques statistiques pour déterminer quels items conserver ou éliminer. Ces techniques sont :

- la technique d'analyse des items de F. Davis (1966) ;
- la technique de Nuttal et Skurnik (1986) et ;
- la technique d'Appel et Kipnis.

a) **Technique d'analyse des items de F. Davis (1966)** : cette méthode, simple et rapide, permet d'évaluer chaque item en calculant simultanément son indice de difficulté et de discrimination, en se concentrant sur les groupes extrêmes (27% supérieurs et 27% inférieurs). La formule considère le nombre de réponses correctes, incorrectes ou non atteintes dans ces groupes. Les indices de facilité (F) et de discrimination (IDs) sont extraits des tables spécifiques, et un item est retenu s'il présente un indice de difficulté entre .20 et .80, et une discrimination supérieur à .20.

b) **Technique de Nuttal et Skurnik** : après analyse, cette technique permet d'interpréter l'indice de facilité (F) de la manière ci-après :

- $F \leq .10$: item trop difficile, à éliminer (IE)
- $F \geq .90$: item trop facile, à éliminer (IE)
- $.75 \leq F < .90$; item très facile, à utiliser avec prudence (IP)
- $.10 < F \leq .25$: item très difficile, à utiliser avec prudence (IP)
- $.25 < F < .75$: item à utiliser avec confiance (IC)

- c) **Technique d'Appel et Kipnis** : elle sert à interpréter l'indice de discrimination (IDs) en fonction de la taille de l'échantillon. comme Nzokana Te Nzabema F. et al., (2016 :5) proposent dans le tableau ci-dessous :

Tableau 3. Tableau d'interprétation d'indice de discrimination d'Appel et Kipnis

N	Item éliminé (IE)	Item à utiliser avec Prudence (IP)	Item à utiliser avec confiance (IC)
50	.22	.22 – .31	.32
51 – 75	.20	.21 – .27	.28
76 – 100	.19	.20 – .24	.25
101 – 200	.18	.19 – .22	.23
151 – 200	.16	.17 – .21	.22

Source: Nzokana Te Nzabema F. et al., (2016:5)

La lecture de ce tableau est basée sur la taille de l'échantillon, c'est-à-dire, lors de l'interprétation, l'on doit chercher à localiser au sein du tableau (Première colonne N) la taille de son échantillon, puis interpréter (IE, IP, ou IC). Tout en restant strictement dans la rangée où la taille de l'échantillon est localisé. (Gbaka Ndaya P., 2024: 25).

Dans la présente, la partie mise en surbrillance est celle qui nous intéresse, car l'échantillon de travail est de 100 sujets.

3 Résultats

Les résultats de cette étude démontrent clairement le déroulement de l'essai des items. Les résultats comprennent une présentation détaillée des notes brutes, une constitution de deux groupes extrêmes, une analyse de la courbe des résultats, une estimation du temps de passation, ainsi qu'une analyse approfondie des items qui aboutit au reclassement des items retenus.

3.1. Essai des items

Les items que nous avons élaborés constituent une version provisoire du test. Il est crucial de les soumettre à l'épreuve auprès d'un petit échantillon représentatif de la population cible. Cette étape, désignée sous le terme d'essai des items, offre l'opportunité de recueillir des indications précieuses sur plusieurs aspects. D'une part, elle permet d'évaluer la difficulté de chaque item : ceux qui apparaissent trop faciles ou trop difficiles peuvent être éliminés ou modifiés afin d'affiner l'échelle. D'autre part, elle fournit des retours sur la pertinence et la clarté des consignes, ainsi que sur leur compréhension par les répondants. Enfin, cet essai offre des informations essentielles concernant le temps nécessaire à l'administration du test, la durée de la correction, ainsi que les éventuelles difficultés liées à la notation.

Nous avons réalisé cet essai auprès d'un panel d'étudiants issus de la première année de Licence-Master-Doctorat en Psychologie, GAISF (deux vacations jour et soir) et Orientation Scalair et Professionnelle (deux vacations jour et soir) de l'Université Pédagogique Nationale (UPN).

Les résultats bruts obtenus par chaque participant sont présentés dans le tableau ci-dessous. Cependant, il est important de souligner qu'un score brut, pris isolément, ne peut être interprété de manière significative. En effet, il est intrinsèquement influencé par la difficulté des items et par la diversité des compétences des sujets. La démarche la plus éclairée consiste donc à établir un barème de référence, permettant de situer les performances individuelles dans une perspective comparative. (Gbaka Ndaya P., 2024 : 29).

Thème 1. Notes Brutes obtenues par les sujets

Les notes brutes attribuées aux sujets ayant participé à l'administration de test sont présentées dans le tableau ci-dessous :

Tableau 4. Notes brutes par sujet

Sujets	Notes Brutes	Sujets	Notes Brutes
1	47	51	33
2	46	52	33
3	45	53	32
4	44	54	32
5	43	55	31
6	43	56	31
7	43	57	31
8	43	58	31
9	43	59	31
10	43	60	31
11	42	61	30
12	42	62	30
13	41	63	30
14	41	64	29
15	41	65	29
16	40	66	29
17	40	67	29
18	40	68	28
19	40	69	28
20	39	70	27
21	39	71	26
22	39	72	26
23	38	73	26
24	38	74	25
25	38	75	25
26	38	76	25
27	38	77	25
28	38	78	25
29	37	79	25
30	37	80	24
31	36	81	24
32	36	82	23
33	36	83	23
34	36	84	22
35	36	85	21
36	36	86	20
37	36	87	20
38	35	88	19
39	35	89	18
40	35	90	18
41	34	91	17
42	34	92	17
43	34	93	17
44	34	94	16
45	34	95	16
46	33	96	13
47	33	97	13
48	33	98	12
49	33	99	10
50	33	100	7
		Maximum	47
		Minimum	7
		Σ	3121
		Moyenne	31,21
		Ecart-type	8,97054101 \cong 8,97

Source: Gbaka Ndaya P., mémoire de licence (2024: 31)

Les données de ce tableau renseignent sur les notes brutes des étudiants. Il s'observe que la note brute la plus élevée est de 47, tandis que la plus basse est de 7. La moyenne des notes s'établit à 31, 21, avec un écart-type de 8,97, ce qui indique une certaine dispersion des résultats.

Thème 2. Constitution de deux groupes extrêmes

Le tableau suivant présente la répartition des sujets en deux groupes extrêmes, représentant chacun 27% des participants supérieurs et inférieurs pour l'analyse des items.

Tableau 5. Groupe extrêmes de notes brutes

Groupe Supérieur	Notes Brutes	Groupe Inférieur	Notes Brutes
1	47	74	25
2	46	75	25
3	45	76	25
4	44	77	25
5	43	78	25
6	43	79	25
7	43	80	24
8	43	81	24
9	43	82	23
10	43	83	23
11	42	84	22
12	42	85	21
13	41	86	20
14	41	87	20
15	41	88	19
16	40	89	18
17	40	90	18
18	40	91	17
19	40	92	17
20	39	93	17
21	39	94	16
22	39	95	16
23	38	96	13
24	38	97	13
25	38	98	12
26	38	99	10
27	38	100	7

Source: Gbaka Ndaya P., mémoire de licence (2024: 32)

Ce tableau illustre clairement les deux groupes extrêmes sur lesquels se basera l'analyse des items, permettant ainsi une évaluation ciblée des performances.

3.2. Analyse de la courbe des résultats

L'analyse de la forme de la courbe des résultats a été réalisée par le calcul de test du chi-carré pour savoir si les items sont classant ou pas. Le tableau ci-dessous illustre clairement ce point.

Tableau 6. Analyse de la courbe des résultats ($\bar{X}=31,21$ et Ecart-type =8,97)

Classe	Fo	RS	$RS - \frac{x}{s}$	Pc	Fc =Pc.N	Fe	Classes	Fo	Fe	Fo-Fe	(Fo-Fe)²	$\frac{(fo - fe)^2}{fe}$
47-50	1	50,5	2,09	0,982	98,2	3,3	43-50	9	3,3	5,7	32,49	9,85
43-46	8	46,5	1,64	0,949	94,9	6,4	39-42	11	6,4	4,6	21,16	3,306
39-42	11	42,5	1,2	0,885	88,5	10,9	35-38	9	10,9	-1,9	3,61	0,331
35-38	9	38,5	0,76	0,776	77,6	15,4	31-34	12	15,4	-3,4	11,56	0,75
31-34	12	34,5	0,31	0,622	62,2	17,4	27-30	8	17,4	-9,4	88,36	5,078
27-30	8	30,5	-0,13	0,448	44,8	16,7	23-26	8	12,7	-4,7	22,09	1,739
23-26	8	26,5	-0,58	0,281	28,1	12,7	19-22	6	8,3	-2,3	5,29	0,637
19-22	6	22,5	-1,02	0,154	15,4	8,3	15-18	7	4,3	2,7	7,29	1,69
15-18	7	18,5	-1,47	0,071	7,1	4,3	7-14	8	2,8	5,2	27,04	9,66
11-14	3	14,5	-1,91	0,028	2,8	1,9						χ² = 33,041
7-10	5	10,5	-2,36	0,009	0,9	0,9						

Source: Gbaka Ndaya, mémoire de licence (2024:33)

Les résultats de ce tableau sur l'analyse de l'allure de courbe révèle que la valeur du chi-carré calculé (33,041) est supérieure à la valeur critique (18,31) au seuil de signification de 5% avec 10 comme degré de liberté. Cela indique que la distribution ne suit pas une distribution normale selon la courbe de Gauss. Par conséquent, à ce stade, cette épreuve ne permet pas d'obtenir un classement fiable des sujets. Cette situation pourrait être attribuée à la présence de certains items qui ne sont pas idoines.

3.3.Estimation du temps de passation

Le temps total de passation du test pilote de culture générale s'élève à 1764 minutes. En calculant la moyenne arithmétique, nous obtenons un temps de passation d'environ 18 minutes par participant. (Gbaka Ndaya P., 2024: 27).

3.4. Analyse des items

L'analyse des items, utilisant la technique des groupes extrêmes de F. Davis, est facilitée par trois feuilles d'analyse. Les deux premières concernent le groupe supérieur (27% des meilleurs) et le groupe inférieur (27% des faibles).

La feuille du groupe supérieur simplifie le calcul et l'interprétation des pourcentages des forts (PH), tandis que celle du groupe inférieur détaille le processus générant les pourcentages des faibles (PL).

La connaissance de ces pourcentages (PH et PL) permet d'élaborer facilement une troisième feuille d'analyse, un tableau récapitulatif, qui conduit au calcul des indices de facilité (F) et de discrimination (IDs).

Les tableaux 7, 8 et 9 présentent l'essentiel du travail d'analyse des items.

Tableau 7. Feuille d'analyse du groupe supérieur, NH =27%

ITEM	K					Omission	NR _H	R _H	W _H	N _H -NR _H	$\frac{W_H}{K-1}$	$\frac{R_H - \frac{W_H}{K-1}}{N_H - NR_H}$	PH
	A	B	C	D	E								
1	0	0	2	0	25	0	0	25	2	27- 0=27	1	0,89	89
2	0	1	1	0	25	0	0	25	2	27- 0=27	1	0,89	89
3	0	6	2	9	5	5	0	6	16	27- 0=27	4	0,07	7
4	8	0	17	2	0	0	0	17	10	27- 0=27	3	0,52	52
5	26	1	0	0	0	0	0	26	1	27- 0=27	0	0,96	96
6	1	1	0	24	0	1	0	24	2	27- 0=27	1	0,85	85
7	25	0	0	0	1	1	0	25	1	27- 0=27	0	0,93	93
8	24	0	3	0	0	0	0	24	3	27- 0=27	1	0,85	85
9	0	14	7	0	1	5	0	14	8	27- 0=27	2	0,44	44
10	1	25	0	0	0	1	0	1	25	27- 0=27	6	-0,19	-19
11	0	27	0	0	0	0	0	27	0	27- 0=27	0	1	100
12	8	2	17	0	0	0	0	17	10	27- 0=27	3	0,52	52
13	1	0	0	25	0	1	0	25	1	27- 0=27	0	0,93	93
14	26	1	0	0	0	0	0	26	1	27- 0=27	0	0,96	96
15	3	2	20	1	1	0	0	20	7	27- 0=27	2	0,67	67
16	2	0	0	24	0	1	0	24	2	27- 0=27	1	0,85	85
17	0	0	27	0	0	0	0	27	0	27- 0=27	0	1	100
18	0	24	0	0	1	2	0	0	25	27- 0=27	6	-0,22	-22
19	0	0	25	0	1	1	0	25	1	27- 0=27	0	0,93	93
20	0	0	0	27	0	0	0	27	0	27- 0=27	0	1	100
21	26	0	0	0	0	1	0	26	0	27- 0=27	0	0,96	96
22	5	0	5	16	0	1	0	0	26	27- 0=27	7	-0,26	-26
23	1	24	2	0	0	0	0	24	3	27- 0=27	1	0,85	85
24	0	2	5	8	10	2	0	10	15	27- 0=27	4	0,22	22
25	2	4	10	6	1	4	0	10	13	27- 0=27	3	0,26	26
26	9	16	0	1	0	1	0	16	10	27- 0=27	3	0,48	48
27	3	24	0	0	0	0	0	24	3	27- 0=27	1	0,85	85
28	1	0	0	25	0	1	0	25	1	27- 0=27	0	0,93	93
29	0	1	0	0	26	0	0	26	1	27- 0=27	0	0,96	96
30	6	2	1	1	15	2	0	15	10	27- 0=27	3	0,44	44
31	0	19	0	3	4	1	0	19	7	27- 0=27	2	0,63	63
32	27	0	0	0	0	0	0	27	0	27- 0=27	0	1	100
33	2	0	24	1	0	0	0	24	3	27- 0=27	1	0,85	85
34	0	10	14	2	0	1	0	10	16	27- 0=27	4	0,22	22
35	27	0	0	0	0	0	0	27	0	27- 0=27	0	1	100
36	5	2	15	2	3	0	0	15	12	27- 0=27	3	0,44	44
37	3	8	6	9	0	1	0	9	17	27- 0=27	4	0,19	19
38	11	6	0	4	2	4	0	4	19	27- 0=27	5	-0,04	-4
39	1	8	2	13	0	3	0	13	11	27- 0=27	3	0,37	37
40	22	4	0	0	0	1	0	22	4	27- 0=27	1	0,78	78
41	21	5	0	0	0	1	0	21	5	27- 0=27	1	0,74	74
42	0	26	0	0	1	0	0	26	1	27- 0=27	0	0,96	96
43	26	0	0	0	0	1	0	26	0	27- 0=27	0	0,96	96
44	0	6	21	0	0	0	0	21	6	27- 0=27	2	0,70	70
45	9	1	0	16	0	1	0	16	10	27- 0=27	3	0,48	48
46	1	23	2	0	0	1	0	23	3	27- 0=27	1	0,81	81
47	3	20	1	0	1	2	0	20	5	27- 0=27	1	0,70	70
48	5	2	18	0	1	1	0	18	8	27- 0=27	2	0,59	59
49	2	0	1	24	0	0	0	24	3	27- 0=27	1	0,85	85
50	12	1	1	8	0	5	0	12	10	27- 0=27	3	0,33	33
51	4	11	3	2	2	5	0	11	11	27- 0=27	3	0,29	29
52	10	4	4	2	2	5	0	4	18	27- 0=27	5	-0,04	-4
53	2	19	0	0	2	4	0	0	23	27- 0=27	6	-0,22	-22
54	19	0	7	1	0	0	0	19	8	27- 0=27	2	0,63	63
55	0	1	17	5	2	2	0	17	8	27- 0=27	2	0,56	56
56	0	1	0	0	25	1	0	25	1	27- 0=27	0	0,93	93
57	26	0	0	1	0	0	0	26	1	27- 0=27	0	0,96	96
58	1	13	5	0	5	3	0	13	11	27- 0=27	3	0,37	37
59	5	18	0	2	1	1	0	18	8	27- 0=27	2	0,59	59

Source: Gbaka Ndaya, mémoire de licence (2024:35)

Tableau 8. Feuille d'analyse du groupe inférieur, $N_L=27\%$

ITEM	K					Omission	NR _L	R _L	W _L	N _L - NR _L	$\frac{W_L}{K-1}$	$\frac{R_L - \frac{W_L}{K-1}}{N_L - NR_L}$	PL
	A	B	C	D	E								
1	0	0	10	0	15	2	0	15	10	27- 0=27	3	0,44	44
2	0	1	7	0	18	1	0	18	8	27- 0=27	2	0,59	59
3	1	5	5	3	3	10	0	5	12	27- 0=27	3	0,07	7
4	7	1	13	1	0	5	0	13	9	27- 0=27	2	0,41	41
5	9	5	0	0	10	3	0	9	15	27- 0=27	4	0,19	19
6	7	0	0	13	1	6	0	13	8	27- 0=27	2	0,41	41
7	13	3	4	4	1	2	0	13	12	27- 0=27	3	0,37	37
8	17	0	5	1	2	2	0	17	8	27- 0=27	2	0,56	56
9	5	10	5	1	0	6	0	10	6	27- 0=27	2	0,29	29
10	4	17	0	1	1	4	0	4	19	27- 0=27	5	-0,04	-4
11	6	20	1	0	0	0	0	20	7	27- 0=27	2	0,67	67
12	20	1	0	1	1	4	0	0	23	27- 0=27	6	-0,22	-22
13	1	3	1	17	2	3	0	17	7	27- 0=27	2	0,56	56
14	26	0	0	0	0	1	0	26	0	27- 0=27	0	0,96	96
15	5	6	11	0	0	5	0	11	11	27- 0=27	3	0,29	29
16	2	4	1	10	1	9	0	10	8	27- 0=27	2	0,29	29
17	3	0	22	1	0	1	0	22	4	27- 0=27	1	0,78	78
18	5	13	0	1	4	4	0	1	22	27- 0=27	6	-0,19	-19
19	0	0	19	1	1	6	0	19	2	27- 0=27	1	0,67	67
20	0	1	3	22	0	1	0	22	4	27- 0=27	1	0,78	78
21	18	0	3	0	1	5	0	18	4	27- 0=27	1	0,63	63
22	6	0	13	1	3	4	0	3	20	27- 0=27	5	-0,7	-7
23	3	3	13	0	1	7	0	3	17	27- 0=27	4	-0,04	-4
24	1	4	8	0	2	12	0	2	13	27- 0=27	3	-0,04	-4
25	2	4	5	5	2	9	0	5	13	27- 0=27	3	0,07	7
26	10	1	2	0	1	13	0	1	13	27- 0=27	3	-0,07	-7
27	5	12	0	0	0	10	0	12	5	27- 0=27	1	0,41	41
28	2	0	0	15	2	8	0	15	4	27- 0=27	1	0,52	52
29	4	0	0	5	5	13	0	5	9	27- 0=27	2	0,11	11
30	2	6	4	4	4	7	0	4	16	27- 0=27	4	0,04	4
31	5	5	2	6	0	9	0	5	13	27- 0=27	3	0,07	7
32	8	1	1	3	4	10	0	8	9	27- 0=27	2	0,22	22
33	4	1	8	5	3	6	0	8	13	27- 0=27	3	0,19	19
34	3	1	10	0	4	9	0	1	17	27- 0=27	4	-0,11	-11
35	8	7	4	0	0	8	0	8	11	27- 0=27	3	0,19	19
36	6	2	7	1	3	8	0	7	12	27- 0=27	3	0,15	15
37	5	2	2	7	2	9	0	7	11	27- 0=27	3	0,15	15
38	5	7	2	0	3	10	0	0	17	27- 0=27	4	-0,15	-15
39	3	3	2	9	2	8	0	9	10	27- 0=27	3	0,22	22
40	12	1	5	1	2	6	0	12	9	27- 0=27	2	0,37	37
41	15	5	1	1	0	5	0	15	7	27- 0=27	2	0,48	48
42	6	5	1	1	8	6	0	5	16	27- 0=27	4	0,04	4
43	21	0	0	1	0	5	0	21	1	27- 0=27	0	0,78	78
44	4	12	6	0	0	5	0	6	16	27- 0=27	4	0,07	7
45	10	1	1	6	1	8	0	6	13	27- 0=27	3	0,11	11
46	6	5	8	0	2	6	0	5	16	27- 0=27	4	0,04	4
47	6	5	4	2	1	9	0	5	13	27- 0=27	3	0,07	7
48	8	0	4	0	3	12	0	4	11	27- 0=27	3	0,04	4
49	2	0	1	10	3	11	0	10	6	27- 0=27	2	0,29	29
50	3	3	2	2	2	15	0	3	9	27- 0=27	2	0,04	4
51	9	3	0	1	2	12	0	3	12	27- 0=27	3	0	0
52	5	5	3	2	1	11	0	3	13	27- 0=27	3	0	0
53	5	5	4	0	0	13	0	0	14	27- 0=27	4	-0,15	-15
54	10	1	3	0	2	11	0	10	6	27- 0=27	2	0,29	29
55	5	1	6	3	2	10	0	6	11	27- 0=27	3	0,11	11
56	2	2	1	4	5	13	0	5	9	27- 0=27	2	0,11	11
57	4	3	3	4	1	12	0	4	11	27- 0=27	3	0,04	4
58	2	2	7	2	1	13	0	2	12	27- 0=27	3	-0,04	-4
59	3	2	3	7	0	12	0	2	13	27- 0=27	3	-0,04	-4
60	3	3	3	4	3	11	0	3	13	27- 0=27	3	0	0

Source: Gbaka Ndaya, mémoire de licence (2024 :36)

La compréhension de ces deux pourcentages (PH et PL) a permis d'accéder à la table V de F. Davis, où il est possible de déterminer l'indice de facilité (F) et l'indice de discrimination (IDs) pour chaque item de l'épreuve analysée. Cependant, l'utilisation de la table V dépendait parfois des tables II ou III du même auteur.

Tableau 9. Récapitulatif des tableaux NH 27% , NL 27%, et des indices (F) et de (IDs) obtenus

ITEM	K					NR	P	F	IDs
	A	B	C	D	E				
1	0/0	0/0	2/10	0/0	25/15	0/0	89/44/	59	32
2	0/0	1/1	1/7	0/0	25/18	0/0	89/59/	64	25
3	0/1	6/5	2/5	9/3	5/3	0/0	7/7/	20	0
4	8/7	0/1	17/13	2/1	0/0	0/0	52/41/	48	6
5	26/9	1/5	0/0	0/0	0/10	0/0	96/19/	54	61
6	1/7	1/0	0/0	24/13	0/1	0/0	85/41/	57	29
7	25/13	0/3	0/4	0/4	1/1	0/0	93/37/	58	47
8	24/17	0/0	3/5	0/1	0/2	0/0	89/56/	61	21
9	0/5	14/10	7/5	0/1	1/0	0/0	44/29/	43	9
10	1/ 4	25/17	0/0	0/1	0/1	0/0	-19/-4/	25	23
11	0/6	27/20	0/1	0/0	0/0	0/0	100/67/ (98/67/)	70	38
12	8/20	2/1	17/0	0/1	0/1	0/0	52/-22/ (52/2/)	37	51
13	1/1	0/3	0/1	25/17	0/2	0/0	93/56/	64	35
14	26/26	1/0	0/0	0/0	0/0	0/0	96/96/	87	0
15	3/5	2/6	20/11	1/0	1/0	0/0	67/29/	49	24
16	2/2	0/4	0/1	24/10	0/1	0/0	85/29/	54	39
17	0/3	0/0	27/22	0/1	0/0	0/0	100/78/ (98/78/)	75	31
18	0/5	24/13	0/0	0/1	¼	0/0	-22/-19/	32	4
19	0/0	0/0	25/19	0/1	1/1	0/0	93/67/	68	27
20	0/0	0/1	0/3	27/22	0/0	0/0	100/78/ (98/78/)	75	31
21	26/18	0/0	0/3	0/0	0/1	0/0	96/63/	68	34
22	5/6	0/0	5/13	16/1	0/3	0/0	-26/-7/	30	19
23	1/3	24/3	2/13	0/0	0/1	0/0	85/-4/ (85/2/)	47	74
24	0/1	2/4	5/8	8/0	10/2	0/0	22/-4/ (22/2/)	31	25
25	2/2	4/4	10/5	6/5	½	0/0	26/7/	29	23
26	9/10	16/1	0/2	1/0	0/1	0/0	48/-7/ (48/2/)	36	49
27	3/5	24/12	0/0	0/0	0/0	0/0	85/41/	57	29
28	1/ 2	0/0	0/0	25/15	0/2	0/0	93/52/	62	34
29	0/4	1/0	0/0	0/5	26/5	0/0	96/11/	52	70
30	6/2	2/6	1 / 4	1/ 4	15/4	0/0	44/4/	35	39
31	0/5	19/5	0/2	3/6	4/0	0/0	63/7/	41	40
32	27/8	0/1	0/1	0/3	0/4	0/0	100/22/ (98/22/)	55	66
33	2/4	0/1	24/8	1/5	0/3	0/0	85/19/	52	47
34	0/3	10/1	14/10	2/0	0/4	0/0	22/-11/ (22/2/)	25	31
35	27/8	0/7	0/4	0/0	0/0	0/0	100/19/ (98/19/)	55	68
36	5/6	2/2	15/7	2/1	3/3	0/0	44/11/	39	21
37	3/5	8/2	6/2	9/7	0/2	0/0	19/15/	31	4
38	11/5	6/7	0/2	4/0	2/3	0/0	-4/-15/	22	+16
39	1/3	8/3	2/2	13/9	0/2	0/0	37/22/	38	10
40	22/12	4/1	0/5	0/1	0/2	0/0	78/37/	54	27
41	21/15	5/5	0/1	0/1	0/0	0/0	74/48/	56	17
42	0/6	26/5	0/1	0/1	1/8	0/0	96/4/	50	83
43	26/21	0/0	0/0	0/1	0/0	0/0	96/78/	74	24
44	0/4	6/12	21/6	0/0	0/0	0/0	70/7/	44	47
45	9/10	1/1	0/1	16/6	0/1	0/0	48/11/	39	28
46	1/6	23/5	2/8	0/0	0/2	0/0	81/4/	46	61
47	3/6	20/5	1 /4	0/2	1/1	0/0	70/7/	44	47
48	5/8	2/0	18/4	0/0	1/3	0/0	59/4/	40	49
49	2/2	0/0	1/1	24/10	0/3	0/0	85/29/	54	39
50	12/3	1/3	1 / 2	8/2	0/2	0/0	33/4/	32	32
51	4/9	11/3	3/0	2/1	2/2	0/0	29/0/ (29/2/)	29	37
52	10/5	4/5	4/3	2/2	2/1	0/0	-4/0/ (-4/2/)	10	7
53	2/5	19/5	0/4	0/0	2/0	0/0	-22/-15/	31	7
54	19/10	0/1	7/3	1/0	0/2	0/0	63/29/	48	22

55	0/5	1/1	17/6	5/3	2/2	0/0	56/11/	41	32
56	0/2	1 /2	0/1	0/4	25/5	0/0	93/11/	52	66
57	26/4	0/3	0/3	1/4	0/1	0/0	96/4/	50	83
58	1 /2	13/2	5/7	0/2	5/1	0/0	37/-4/ (37/2/)	32	41
59	5/3	18/2	0/3	2/7	1/0	0/0	59/-4/ (59/2/)	40	55
60	5/3	3/3	12/3	1/4	3/3	0/0	33/0/ (33/2/)	31	40

Source: Gbaka Ndaya, mémoire de licence (2024:37)

Après ce tableau 9, l'étape suivante consiste à interpréter chaque indice obtenu, ce qui facilitera ultérieurement la prise de décision.

3.5. Sélection des items

Pour décider quels items doivent être retenus ou écartés, nous nous sommes basé sur deux critères essentiels : indices de difficulté (IDf) et de discrimination (IDs). en d'autres termes, un item sera conservé que s'il présente des valeurs satisfaisantes pour ces deux indicateurs. Plus précisément, l'item retenu doit posséder à la fois un indice de difficulté et un indice de discrimination qui répondent aux seuils établis. Autrement dit, seul l'item qui peut être utilisé avec confiance (IC), c'est-à-dire dont l'indice de facilité se situe entre .20 et .80, et dont l'indice de discrimination est supérieur à .20, sera considéré comme satisfaisant pour l'analyse, conformément aux recommandations de Ronald Verbeke (1970 : 64) et F. B. Davis (1966 :17).

Cette approche rigoureuse nous permettra, dans un premier temps, de réduire de manière significative le nombre d'items à retenir, en veillant à ne maintenir que ceux qui présentent une capacité à distinguer efficacement les différents niveaux de connaissance des répondants. En outre, cette démarche prend en compte non seulement la capacité discriminante de chaque item, mais aussi son niveau de difficulté relatif à l'ensemble, ce qui garantit une sélection équilibrée et pertinente des items. en appliquant ces critères, nous assurons ainsi une amélioration qualitative du contenu de l'épreuve, tout en maintenant une cohérence dans la difficulté et la discrimination des items retenus, ce qui renforcera la fiabilité et la validité de l'évaluation globale.

Tableau 10. Interprétation des indices (IDf et IDs) et élimination des items inadéquats

Item	IDf	Interprétation	IDs	Interprétation	Décision
1	59	IC	32	IC	M
2	64	IC	25	IC	M
3	20	IP	0	IE	R
4	48	IC	6	IE	R
5	54	IC	61	IC	M
6	57	IC	29	IC	M
7	58	IC	47	IC	M
8	61	IC	21	IP	R
9	43	IC	9	IE	R
10	25	IC	23	IP	R
11	70	IC	38	IC	M
12	37	IC	51	IC	M
13	64	IC	35	IC	M
14	87	IP	0	IE	R
15	49	IC	24	IP	R
16	54	IC	39	IC	M
17	75	IP	31	IC	M
18	32	IC	4	IE	R
19	68	IC	27	IC	M
20	75	IP	31	IC	M
21	68	IC	34	IC	M
22	30	IC	19	IE	R
23	47	IC	74	IC	M
24	31	IC	25	IC	M
25	29	IC	23	IP	R
26	36	IC	49	IC	M
27	57	IC	29	IC	M
28	62	IC	34	IC	M
29	52	IC	70	IC	M
30	35	IC	39	IC	M
31	41	IC	40	IC	M
32	55	IC	66	IC	M
33	52	IC	47	IC	M
34	25	IC	31	IC	M
35	55	IC	68	IC	M
36	39	IC	21	IP	R
37	31	IC	4	IE	R
38	22	IP	+16	IE	R
39	38	IC	10	IE	R
40	54	IC	27	IC	M
41	56	IC	17	IE	R
42	50	IC	83	IC	M
43	74	IC	24	IP	R
44	44	IC	47	IC	M
45	39	IC	28	IC	M
46	46	IC	61	IC	M
47	44	IC	47	IC	M
48	40	IC	49	IC	M
49	54	IC	39	IC	M
50	32	IC	32	IC	M
51	29	IC	37	IC	M
52	10	IE	7	IE	R
53	31	IC	7	IE	R
54	48	IC	22	IP	R
55	41	IC	32	IC	M
56	52	IC	66	IC	M
57	50	IC	83	IC	M
58	32	IC	41	IC	M
59	40	IC	55	IC	M
60	31	IC	40	IC	M

Source: Ghaka Ndava. mémoire de licence (2024:39)

Légende : **IC** : Item à utiliser avec Confiance ; **IP** : Item à utiliser avec Prudence ; **IE** : Item à Éliminer ; **M** : Maintien ; **R** : Rejet.

Le tableau 10 présente l'interprétation du sort de chaque item en termes d'indice de facilité (F) et d'indice de discrimination (IDs) ainsi que la décision qui a sanctionné ce travail d'analyse. Conformément aux techniques d'interprétations des deux indices de Nuttal et Skurnik pour l'interprétation de l'indice de facilité des items et d'Appel et Kipnis pour l'indice de discrimination des items. Le traitement statistique s'est soldé par les résultats suivants:

1. Au niveau de la difficulté

- 1 item a été jugé inadéquat et doit donc être éliminé (IE), représentant ainsi 2% de l'ensemble. Il s'agit de l'item 52.
- autres items, à savoir les items 3, 14, 17, 20 et 38, soit 8% du total, doivent être utilisés avec prudence (IP)
- en revanche, 54 items, soit 90%, peuvent être utilisés avec confiance (IC). ces items sont : 1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 57, 58, 59 et 60.

2. Au niveau de la discrimination

- 12 items, représentant 20%, doivent être éliminés. il s'agit des items suivants : 3, 4, 9, 14, 18, 22, 37, 38, 39, 41, 52 et 53.
- 7 items supplémentaires (8, 10, 15, 25, 36, 43 et 54), soit 12% du total, doivent être utilisés avec prudence.
- enfin, 41 items, soit 68% peuvent être utilisés avec confiance. ces items incluent : 1, 2, 5, 6, 7, 11, 12, 13, 16, 17, 19, 20, 21, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 42, 44, 45, 46, 47, 48, 49, 50, 51, 55, 56, 57, 58, 59 et 60.

3. Au niveau décision

- 18 items ont été déclarés inadéquats, représentant 32% du total. les items concernés sont : 3, 4, 8, 9, 10, 14, 15, 18, 22, 25, 36, 37, 38, 39, 41, 43, 52, 53 et 54. le chi-carré calculé est de 22, chi-carré calculé (33,041) est supérieure à la valeur critique (18,31) au seuil de signification de 5% avec 10 comme degré de liberté. cela confirme la présence de ces items inadéquats.
- 41 items suivants : 1, 2, 5, 6, 7, 11, 12, 13, 16, 17, 19, 20, 21, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 42, 44, 45, 46, 47, 48, 49, 50, 51, 55, 56, 57, 58, 59 et 60 ont été retenus et peuvent être utilisés en toute confiance (IC). ces items se révèlent idoines grâce à leurs indices de discrimination et de facilité jugés satisfaisants dans les deux cas. **Par conséquent, notre hypothèse de travail a été confirmée.**

Tableau 11. Présentation des items retenus

Item	IDf	Interprétation	IDs	Interprétation	Décision
1	59	IC	32	IC	M
2	64	IC	25	IC	M
5	54	IC	61	IC	M
6	57	IC	29	IC	M
7	58	IC	47	IC	M
11	70	IC	38	IC	M
12	37	IC	51	IC	M
13	64	IC	35	IC	M
16	54	IC	39	IC	M
17	75	IP	31	IC	M
19	68	IC	27	IC	M
20	75	IP	31	IC	M
21	68	IC	34	IC	M
23	47	IC	74	IC	M
24	31	IC	25	IC	M
26	36	IC	49	IC	M
27	57	IC	29	IC	M
28	62	IC	34	IC	M
29	52	IC	70	IC	M
30	35	IC	39	IC	M
31	41	IC	40	IC	M
32	55	IC	66	IC	M
33	52	IC	47	IC	M
34	25	IC	31	IC	M
35	55	IC	68	IC	M
40	54	IC	27	IC	M
42	50	IC	83	IC	M
44	44	IC	47	IC	M
45	39	IC	28	IC	M
46	46	IC	61	IC	M
47	44	IC	47	IC	M
48	40	IC	49	IC	M
49	54	IC	39	IC	M
50	32	IC	32	IC	M
51	29	IC	37	IC	M
55	41	IC	32	IC	M
56	52	IC	66	IC	M
57	50	IC	83	IC	M
58	32	IC	41	IC	M
59	40	IC	55	IC	M
60	31	IC	40	IC	M

Source: Gbaka Ndaya, mémoire de licence (2024: 42)

Le tableau ci-dessus présente tous les items retenus après leur analyse. Il est important de signaler que les chiffres dans la première colonne (Items) ne suivent pas l'ordre numérique croissant, mais correspondent aux numéros attribués à ces items avant leur analyse.

Tableau 12. Présentation du reclassement des items retenus

Item N° Initial	Item N° Actuel	IDf	Interprétation	IDs	Interprétation	Décision
34	1	25	IC	31	IC	M
51	2	29	IC	37	IC	M
24	3	31	IC	25	IC	M
60	4	31	IC	40	IC	M
50	5	32	IC	32	IC	M
58	6	32	IC	41	IC	M
30	7	35	IC	39	IC	M
26	8	36	IC	49	IC	M
12	9	37	IC	51	IC	M
45	10	39	IC	28	IC	M
48	11	40	IC	49	IC	M
59	12	40	IC	55	IC	M
31	13	41	IC	40	IC	M
55	14	41	IC	32	IC	M
44	15	44	IC	47	IC	M
47	16	44	IC	47	IC	M
46	17	46	IC	61	IC	M
23	18	47	IC	74	IC	M
42	19	50	IC	83	IC	M
57	20	50	IC	83	IC	M
29	21	52	IC	70	IC	M
33	22	52	IC	47	IC	M
56	23	52	IC	66	IC	M
5	24	54	IC	61	IC	M
16	25	54	IC	39	IC	M
40	26	54	IC	27	IC	M
49	27	54	IC	39	IC	M
32	28	55	IC	66	IC	M
35	29	55	IC	68	IC	M
6	30	57	IC	29	IC	M
27	31	57	IC	29	IC	M
7	32	58	IC	47	IC	M
1	33	59	IC	32	IC	M
28	34	62	IC	34	IC	M
2	35	64	IC	25	IC	M
13	36	64	IC	35	IC	M
19	37	68	IC	27	IC	M
21	38	68	IC	34	IC	M
11	39	70	IC	38	IC	M
17	40	75	IP	31	IC	M
20	41	75	IP	31	IC	M

Source: Gbaka Ndaya, mémoire de licence (2024: 44)

Les données de ce tableau renseignent sur le reclassement des items idoines. Il s'observe que ces items ont été reclassés afin d'optimiser les performances des sujets, car ils constitueront la première version expérimentale de notre épreuve. Le reclassement se fait selon un ordre croissant de difficulté, commençant par les items très faciles, suivis des items faciles, puis des items difficiles et très difficiles.

4. Discussion de résultats

Le test psychologique est un outil de prédilection pour les conseillers d'orientation et les psychologues, tant pour le testing que pour l'évaluation scolaire et psychologique des individus. Cependant, son efficacité repose sur sa capacité à s'adapter aux réalités spatiotemporelles et linguistique. La mauvaise utilisation des tests peut causer des torts considérables aux usagers et aux intervenants concernés par les décisions fondées sur ces tests. Sarrazin G., (2003)

Dans le contexte de la République Démocratique du Congo (RDC), le manque d'instruments de mesure approprié susceptible d'établir un diagnostic fiable des compétences ou aptitudes et des besoins psychologiques des individus constitue un véritable défi (Ministère de l'Education Nationale, 2015). Selon une étude publiée dans le Journal of African Psychology, les tests psychologiques utilisés en RDC sont souvent conçus et adaptés à des cultures et des contextes différents, ce qui peut entraîner des biais et des erreurs de diagnostic (Makura et al., 2018).

En outre, la recherche a montré que la fiabilité et la validité des tests psychologiques dépendant de plusieurs facteurs, notamment de la qualité de la conception et de la validation des instruments, de la formation et de l'expérience du psychologue qui administre le test, ainsi que de la sensibilité et de la spécificité des résultats (Américan Psychological Association, 2001). Un besoin majeur se fait sentir en RDC, car l'insuffisance des spécialistes amène à la confusion persistante entre épreuves, concours et tests, notamment dans les milieux scolaires et universitaires, illustre cette problématique. En effet, le Ministère de l'Enseignement Primaire, Secondaire et Technique (MEPST) en RDC semble assimiler certaines épreuves telles que le TENAFEP⁵, le TENASOP⁶, l'Examen d'Etat ou encore les concours d'entrée dans les établissements scolaires ou universitaires à des tests, alors que, comme l'indique Azia Dimbu (2024), toute épreuve n'est pas nécessairement un test. Un véritable test doit impérativement respecter les critères fondamentaux de validité, de fidélité et de sensibilité qui sont les fondamentales métrologiques du test.

L'objectif principal de cette étude est de concevoir un outil original, adapté au contexte congolais, capable d'évaluer les connaissances culturelles des étudiants. Les questions suivantes ont été soulevées : les items du test pilote sont-ils idoines ? Combien d'entre eux sont inadéquats ? Quel est le moment plus propice pour administrer ce test afin d'obtenir des résultats significatifs ? Cette version du test nécessite-t-elle des ajustements avant sa mise en œuvre définitive ? Une réponse provisoire émise indiquant que le test pilote de culture générale d'inspiration congolaise est majoritairement constitué d'items idoines pour l'évaluation des étudiants. Pour confirmer ou infirmer cette hypothèse, une enquête basée sur le test a été menée afin de collecter des données. De plus, nous avons appliqué les techniques statistiques de Davis, Nuttal, Skurnik, Appel et Kipnis, pour analyser ces résultats et sélectionner les meilleurs items.

L'épreuve a été soumise à un échantillon de 100 étudiants de première année de Licence-Master-Doctorat (L1 LMD) inscrits en Psychologie, GAISF⁷ (jour et soir) et OSP⁸ (jour et Soir) à l'Université Pédagogique Nationale (UPN). Il s'agit d'un échantillon occasionnel, représentant une partie de la population étudiante, choisi pour des raisons d'accessibilité. ce choix ne repose pas sur une logique scientifique stricte, mais plutôt sur la disponibilité des sujets, conformément aux principes de Quivy et Van Campenhout (2006). Par conséquent, les résultats obtenus ne peuvent pas être généralisés à l'ensemble de la population cible.

⁵ Test National de Fin d'Etudes Primaires(TENAFEP), est un examen national important en RDC pour les élèves finalistes de la 6^{ème} année primaire, sanctionnant la fin de leur cycle et servant d'orientation pour le secondaire souvent remplacé ou appelé actuellement ENAFEP (Examen National de Fin d'Etudes Primaires).

⁶ Test National de Sélection et d'Orientation Scolaire et Professionnelle (TENASOP), est l'examen annuel pour les élèves de 8^{ème} année de l'éducation de base (équivalent de la 9^{ème} année), afin de déterminer leur orientation future vers les humanités ou la formation professionnelle, basé sur leurs aptitudes psychologiques et leurs résultats, marquant la fin du cycle primaire.

⁷ Gestion et Administration des Institutions Scolaires et des Formations.

⁸ Orientation Scolaire et Professionnelle.

Après analyse statistique, les résultats se synthétisent selon quatre axes principaux ci-dessous :

1. Niveau de la difficulté

Un item (item 52) a été éliminé, représentant 2% du total. 5 items (3, 14, 17, 20 et 38) soit 8%, doivent être utilisés avec prudence, tandis que 54 items (1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 57, 58, 59, 60) peuvent être utilisés avec confiance, ce qui représente environ 90%. Selon Davis (1966) la difficulté d'un item se définit par le pourcentage de répondants corrects ou par la proportion de sujets connaissant la réponse.

2. Niveau de la discrimination

Douze items (3, 4, 9, 14, 18, 22, 37, 38, 39, 41, 52, 53), ont été éliminés, soit 20%. 7 items autres (8, 10, 15, 25, 36, 43, 54), soit 12%, doivent être utilisés avec prudence. Enfin, 41 items (1, 2, 5, 6, 7, 11, 12, 13, 16, 17, 19, 20, 21, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 42, 44, 45, 46, 47, 48, 49, 50, 51, 55, 56, 57, 58, 59, 60) peuvent être utilisés avec confiance. La capacité de discrimination d'un item, selon Findley, (1956), est sa faculté à différencier les participants performants des moins performants, en se basant sur la différence d'indice de difficulté entre groupes à performance élevée et faible. ce critère est essentiel pour assurer que chaque item contribue à différencier efficacement les niveaux de compétence. Ebel et Frisbie (1991).

3. Décisions prises

Deux décisions ont été formulées : la première consiste à écarter 18 items (3, 4, 8, 9, 10, 14, 15, 18, 22, 25, 36, 37, 38, 39, 41, 43, 52, 53, 54), représentant 32% des items inadéquats. la seconde consiste à conserver 41 items (1, 2, 5, 6, 7, 11, 12, 13, 16, 17, 19, 20, 21, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 42, 44, 45, 46, 47, 48, 49, 50, 51, 55, 56, 57, 58, 59, 60), soit 68% des items jugés idoines. Ces résultats confirment l'hypothèse selon laquelle un bon item doit présenter un indice de difficulté compris entre .20 et .80, conformément aux normes proposées par Davis (1966), Azia Dimbu (2024), Verbeke (1970) et Laveault et Grégoire (2008). Ce critère est largement accepté dans la littérature pour assurer la qualité des items dans un test. Hambleton et Swaminathan (1985).

4. Temps d'administration

Le temps moyen de passation du test s'est révélé être de 18 minutes, ce qui indique un outil relativement rapide à administrer dans le contexte universitaire congolais. La durée d'un test d'efficiency mentale varie généralement entre 10 et 30 minutes selon les exemples, tels que l'Inventaire d'intérêts professionnels de Rothwell-Miller (IRMR, 2011), le test de Strong (1927) ou celui de John Holland (RIASEC⁹, 12 minutes) illustrent cette tendance (Holland, 1997).

⁹ RIASEC, est un test d'orientation professionnelle basé sur la théorie de John Holland, qui classe les personnalités en six types notamment ; Realiste, Investigateur, Artiste, Social, Entreprenant et Conventionnel.

Conclusion

Au terme de cette étude, il apparaît que le test pilote de culture générale, conçu dans une vision d'adaptation au contexte congolais, montre une majorité d'items idoines pour l'évaluation des connaissances des étudiants. La méthode d'analyse appliquée, combinant la technique de F. Davis et les techniques de Nuttal, Skurnik, Appel et Kipnis, a permis de distinguer nettement les items efficaces de ceux nécessitant une révision.

Le temps moyen de passation, estimé à 18 minutes, témoigne de la simplicité et de la rapidité de l'outil, ce qui est un atout dans le contexte universitaire.

L'analyse du niveau de difficulté révèle que la majorité des items (90%) sont jugés utilisables avec confiance, même si un petit nombre doit être éliminé ou utilisé avec prudence. Sur le plan de la discrimination, 68% des items se sont avérés adéquats pour différencier les étudiants selon leur capacité, ce qui conforte la validité partielle du test. Cependant, un certain nombre d'items (20%) ont été identifiés comme inadéquats ou peu discriminants, nécessitant leur retrait ou leur modification pour renforcer la fiabilité de l'épreuve.

Les résultats confirment également que le test, dans sa version actuelle, constitue un prototype encore perfectible. La durée d'administration, bien qu'adéquate, doit être affinée pour garantir une meilleure stabilité des résultats. Par ailleurs, le traitement statistique souligne la nécessité d'un ajustement des items pour optimiser leur capacité à mesurer de manière précise et équilibrée la culture générale.

Il est important de souligner que cette étape constitue une avancée significative dans la création d'outils d'évaluation culturellement adaptés. La pertinence des items, bien qu'encourageante, doit être renforcée par des améliorations pour assurer une validité optimale. Les limites de l'étude, notamment la taille de l'échantillon et le contexte spécifique de passation, doivent également être prises en compte dans le cadre de futures recherches.

Ces résultats offrent des pistes concrètes pour l'amélioration du test, qui pourrait devenir un instrument fiable pour l'évaluation de la culture générale dans le système universitaire congolais. La poursuite de ce travail permettra d'affiner l'outil, en intégrant de nouveaux items et en ajustant ceux existants, afin de garantir une évaluation juste, complète et adaptée aux réalités spatiotemporelles.

REFERENCES

- American Psychological Association. (2001). *Publication Manual of the American Psychological Association*. (5th Ed.). Washington : Auteur.
- Azia Dimbu, F. (2024). *Expérimentation d'un test psychologique d'efficacité mentale : élucidation de grandes articulations*. Akofena. D.O.I :<https://doi.org/10.48734/akofena>
- Azia Dimbu, F., Kodila Tedika, O. et Kimboko Mpesi, J. (2019). *Normes de présentation d'un travail scientifique*. Paris :L'Harmattan.
- Azia Dimbu, F., Nzokana Te Nzabema, F. et Bahuidi Salabiaku V. (2016). *Essai des items d'une épreuve de signification verbale à Kinshasa*. Presses de l'Université Pédagogique Nationale, n°066a (2016), Janvier-Mars 2016, p.211-230.
- Chartier, P. et Loarer, E. (2008). *Evaluer l'intelligence logique. Approche cognitive et dynamique*. Paris : Dunod.
- Davis, F. B. (1966). *Analyse des items*. Louvain : Nauwelaerts.
- Ebel, R. L. et Frisbie, D. A. (1991). *Essential of educational measurement*. Englewood Cliffs, NJ : Prentice Hall.
- Findley, W. G. (1956). A rationale for evaluation of item discrimination statistics. *Educational and Psychological Measurement*, 16, 175-180.
- Gbaka Ndaya, P.(2024). *Essai des items d'un test pilote de culture générale aux étudiants de l'Université Pédagogique Nationale*. Mémoire de licence, Kinshasa: UPN-PSO.
- Hambleton, R. K., et Swaminathan, H. (1985). *Item response theory : principles and Applications*. Dordrecht, the Netherlands : Kluwer.

- Laveault, D. et Grégoire, J. (2014). *Introduction aux théories des tests : en psychologie et en sciences de l'éducation*. 3^{ème} édition. Paris : De Boeck Supérieur.
- Lievens, F. et Patterson, F. (2011). *The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection*. Journal of Applied Psychology, vol.96, n°5, P.927-940.
- Mabiala Seda et Banzikila Sile. (1988). *Etalonnage des tests : utilité et pratique*. Scientia,2,(2),167-184.
- Makura, E., Mpongo, S., et Mumbwa, J. (2018). *L'utilisation des tests psychologiques en République Démocratique du Congo : enjeux et défis*. Journal of African Psychology, 32(1), 1-12.
- Piaget, J. (1972). *Psychologie et épistémologie*. Paris : Denoël.
- Piaget, J., & Inhelder, B. (1966). *La psychologie de l'enfant*. Paris : Presses Universitaires de France.
- Quivy, R. et Van Campenhoudt, L. (2006). *Manuel de recherche en sciences sociales*. Paris : Dunod.
- Rathus S.A. (1991). *Psychologie générale*, 2^{ème} édition. Etudes Vivantes, Mor Traduction de Lyna Lepage.
- Sarrazin, G. (2003). *Normes de pratique du testing en psychologie et en éducation*. Montréal : Institut de Recherches Psychologiques.
- Verbeke, R. (1970). *La compréhension du vocabulaire dans l'apprentissage d'une langue étrangère*. Kinshasa : Publications de l'Université Lovanium.