



Qualités métrologiques du test de maturité mentale Columbia en milieu algérien

الخصائص المترولوجية (القياسية) لاختبار النضج العقلي
للكولومبيا في البيئة الجزائرية

Metrological qualities in Columbia in the Algerian environment

Dr. Sadjia Makhlouf-Bentounes

Pr. Fatima Arar

Pr. Fatima Zinet

Université Alger 2

Date de soumission: 23-10-2019-Date d'acceptation: 21-11-2019-Date de publication: 08-12-2020

ملخص

يهدف هذا المقال إلى تقديم بعض جوانب عمل فرقة البحث، التي تدرس موضوع "تكييف اختبار النضج العقلي للكولومبيا في الوسط الجزائري" وهو اختبار أنشأ من طرف ب. بجمستر وآخرون (1947-1972) بهدف كشف الذكاء التصنيفي وهو اختبار غير لفظي معروف، و كثير التداول في الميدان العيادي، غير أن استعماله يتطلب تكييفاً في البيئة الجزائرية. و من أجل جعل هذا الإختبار ممكن التطبيق في الوسط الجزائري، قمنا في المرحلة الأولى من هذا البحث بدراسة ثباته و صدقه ذلك من خلال تقييم قدرته على التمييز بين الأفراد؛ واستمرارية نتائجه عبر الزمن، وأيضاً صدقه فيقياس ما وضع من أجله. وبناءً على ما سبق يتضح أننا سوف نركز في هذه المقالة على الدراسة الإجرائية لمفهوم الصدق والثبات، وتوظيفهما في الدراسة الميدانية لهذا البحث، تحت ضوء المنهجية المتبعة في دراسة إشكالية تكييف مقياس الكولومبيا في الوسط الجزائري.

الكلمات الدالة: مقياس النضج العقلي الكولومبيا؛ ثبات المقياس؛ صدق المقياس؛ تكييف المقياس في البيئة الجزائرية.

Résumé

Cet article a pour objectif de présenter certains aspects du travail de notre équipe de recherche qui porte sur «*L'adaptation du test de maturité mentale du Colombia en milieu algérien*». Le test de B. Burgeimeister et collaborateurs (1947-1972) dont il est question, est un test d'intelligence qui explore particulièrement l'activité catégorielle et classificatoire (Dague, 1965). C'est aussi un test non verbal, notoire dont l'utilisation en milieu clinique algérien est très courante. Il n'en demeure pas moins que son adaptation en Algérie reste à faire. C'est dans l'objectif de rendre ce test praticable en

milieu algérien que nous avons entrepris d'expérimenter les concepts de fidélité et de validité du test de maturité mentale du Columbia, autrement dit évaluer:

- sa sensibilité (sa capacité à différencier entre les individus),
- sa stabilité dans le temps
- sa pertinence (s'il mesure bien ce qu'il prétend mesurer).

Cet article mettra, donc, l'accent sur le travail d'opérationnalisation de ces concepts et leur expérimentation sur le terrain.

Mots-clés: test de maturité mentale du Columbia; fidélité du test; validité du test; adaptation en milieu Algérien.

Abstract

This paper aims to present some aspects of the work of our research team which focuses on "Adaptation of mental maturity test of Columbia in Algerian environment". The test of B. Burgeimeister et al. (1947-1972) in question is an intelligence test that particularly explores categorical and classificatory activity (Dague, 1965). It's also a non-verbal test, notorious test whose use in an Algerian clinical setting is very common. Nevertheless, its adaptation in Algeria remains to be done.

It is for this purpose to make this test practicable in the Algerian environment that we undertook to test the concepts of fidelity and validity of the Columbia Mental Maturity test; in other words, evaluate:

- Its sensitivity (its ability to differentiate between individuals);
- Its stability over time;
- Its relevance (if it measures what it claims to measure).

The presentation of this paper will focus on the operationalization of these concepts and their field experimentation.

Keywords: Columbia mental maturity test; test fidelity; test validity; adaptation in Algerian environment.

Introduction

Le présent article expose une partie des résultats d'une étude réalisée dans le cadre d'un projet de recherche CNEPRU¹ au niveau du laboratoire de psychologie clinique et métrique portant sur l'«Adaptation du test de maturité mentale du Columbia en milieu algérien.».

¹ - sous la direction du Pr. ARAR F



La démarche suivie consiste à présenter, dans un premier temps, des données sur ce test pour monter son importance méthodologique, exposer, ensuite, quelques expériences de son application dans le monde et présenter, dans un troisième temps, les résultats de l'étude de l'adaptation de cet outil en milieu algérien.

1. Etat de connaissances sur le teste Columbia

1.1 le test de Maturité Mentale du Columbia

C'est un test d'intelligence qui a été mis au point entre 1947 et 1959 par Bessie Burgemeister et coll. à l'université de Columbia, puis revu en 1972. Il s'agit d'une épreuve non verbale, se composant d'une centaine de planches, où figurent des dessins géométriques ou des objets «qui exige un effort de discrimination et consiste à reconnaître le dessin qui ne va pas avec les autres et à découvrir le principe d'organisation des dessins. Ce test sollicite surtout la pensée conceptuelle ou catégorielle ...» (Dague et coll., 1965). Il a été élaboré pour répondre à l'examen psychologique des enfants infirmes moteurs cérébraux, mais il a été progressivement utilisé avec des enfants bien portants. Sa facilité d'utilisation et surtout sa non saturation d'éléments culturels nous a amené à décider de l'adapter à la population algérienne.

1.2 Définitions des concepts

Selon Pichot (1999) «un test est une situation expérimentale standardisée servant de stimulus à un comportement. Ce comportement est évalué par comparaison statistique avec celui d'autres individus placés dans la même situation, permettant ainsi, de classer le sujet examiné, soit quantitativement soit typologiquement. Ainsi, avant d'utiliser un test, il faut donc l'évaluer, afin de connaître ses qualités psychométriques propres (sensibilité, fidélité, validité prédictive)...». Ce sont ces qualités psychométriques qui nous intéressent dans ce travail, que nous allons d'abord définir puis nous attacher à montrer comment ils ont été appliqués au test du Columbia.



1.3 La sensibilité

La sensibilité d'un test représente le pouvoir classant du test, sa capacité à discriminer les éléments testés lorsqu'ils diffèrent, et sa capacité à regrouper ces éléments lorsqu'ils convergent ou sont similaires.

Pour C. Guillevic et S.Vautier (1998) «Un instrument est dit sensible s'il permet de différencier un nombre suffisant de degrés d'une même variable (...). Il doit permettre également de différencier correctement deux individus sur un trait ou une conduite mesuré (Par exemple l'autorité, l'agressivité etc...) et dont la puissance discriminative est fonction des objectifs poursuivis par le chercheur qui peut ranger ses variables sous forme de classes dichotomiques, ou à l'aide de gradations plus fines en plusieurs niveaux». Un test n'est pas assez sensible lorsqu'il regroupe trop d'éléments testés qui diffèrent visiblement.

Une méthode simple pour éprouver la sensibilité d'un test consiste à le faire passer à un nombre suffisamment important de sujets. A partir des résultats obtenus, on peut construire un histogramme par lequel on va pouvoir vérifier à première vue, la distribution des fréquences. Si le test est sensible, les scores devraient être distribués normalement (courbe en cloche suivant la distribution de la loi normale; comme les échelles en quantiles en 5, 7, 11 classes, en déciles, en percentiles etc...).

1.4 La fidélité

La fidélité d'un test renvoie au degré avec lequel une mesure est obtenue avec peu d'erreur. Il y a dans cette idée, la notion de stabilité dans le temps (une mesure réalisée à un moment donné doit fournir la même valeur à un autre moment pour lequel les mêmes conditions sont réunies); elle exige de la précision pour déterminer la valeur sans trop d'erreur.

Selon M. Huteau et J. Lautrey; (1999) «La théorie de la fidélité postule qu'il est possible de décomposer toute mesure observée en 2 parties indépendantes: une mesure vraie non observable à laquelle vient s'ajouter une erreur aléatoire (...) La



moyenne d'une série de mesures observées sera donc une estimation de la mesure vraie...» (p. 98- 99)

«Pour évaluer la fidélité d'une opération de mesure, il faut déterminer la proportion de la variation observée attribuables aux influences transitoires (...). L'instrument est d'autant plus fidèle que cette influence est faible ...» (Belanger et coll., 1988, p.222).

M. Huteau et J. Lautrey déclarent que Le même test étant appliqué à deux reprises (méthode test-retest), le coefficient de corrélation entre les scores obtenus à chaque passation est le coefficient de stabilité et de constance (...)et le plus souvent on calcule directement le coefficient de fidélité sous cette forme et on montre que ce coefficient (de corrélation) est égal entre la première série de mesures et sa répétition...». Il existe, selon plusieurs procédures qui permettent de quantifier cette stabilité; voici les plus classiques (Reuchlin, 1992; Pichot, 1971)

1.4.1 Fidélité temporelle (ou fidélité test-retest et/ou indice d'équivalence)

Le test doit fournir à peu près le même résultat s'il est appliqué deux fois de suite au même sujet au cours de deux passations. Pour calculer un indice de la stabilité (concernant la fidélité pour une mesure individuelle), on utilise le coefficient de Bravais-Pearson. Celui-ci donne une estimation du degré de liaison linéaire entre deux variables numériques recueillie sur un échantillon (les deux variables étant, dans ce cas, les deux scores du test et du retest).

Ce coefficient, noté, servira de coefficient de stabilité (du score individuel aux deux tests. Il est généralement normalisé et compris entre -1 et 1.

Dans le cadre de l'indice d'équivalence « on applique deux épreuves parallèles à un même groupe de sujet (exemple la corrélation entre deux tests proches tel que le D48 et le PM38...» (Guillevic et Vautier (1998), ceci afin de voir s'il n'y a pas d'éléments perturbateurs.



1.4.2 Fidélité interne: (indice d'homogénéité, ou Split-Half)

Le test doit également comporter des items cohérents, et en pratique, «on partage l'épreuve en deux parties équivalentes... on calculera la corrélation entre les résultats (des deux parties) et on vérifie par-là que le test se comporte de manière homogène». (Guillevic; Vautier, 1998) ce qui veut dire qu'ils renvoient au même type de difficultés.

La fidélité interne se calcule à l'aide de différents tests statistiques: le R de Bravais Pearson, le T de Student, le coefficient alpha de Cronbach etc...Un test ayant un coefficient alpha compris entre 0,7 et 1 est considéré comme fidèle.

1.5 La validité

La validité d'un test est sa qualité à mesurer effectivement ce qu'il est censé mesurer. Un test psychologique valide et mesure de façon correcte la dimension psychologique pour la mesure de laquelle il a été conçu et donne des résultats consistants et fiables lorsqu'il est comparé à d'autres instruments de mesure considérés valides et permettant de prédire le comportement.

1.5.1. Validité prédictive

«La valeur d'une recherche dépend nécessairement de l'instrument de mesure utilisé, il faut donc se demander ce que mesure cet instrument, s'agit-il bien de la caractéristique visée (...). La validité est estimée d'après le degré de compatibilité des résultats avec d'autres faits pertinents comme on peut l'apprécier en fonction de son efficacité.»(Belanger et coll., 1988, p.223)

1.5.2. Validité de contenu et validité de construction

«Un test est dit valide lorsqu'il permet d'atteindre de manière satisfaisante les objectifs que le constructeur a choisis. On dira que le test a une bonne validité empirique lorsqu'il apporte une aide effective dans ce genre de situation. On dira qu'il a une bonne validité de contenu, si le test constitue un échantillon représentatif d'un ensemble de tâches ou de conduites. On dira que le test possède une bonne validité théorique s'il permet d'émettre des hypothèses sur le fonctionnement et l'élaboration des réponses du sujet. (Exemples: une échelle



d'intelligence basée sur les théories de l'intelligence)». (Guillevic; Vautier, 1998).

«On caractérise souvent les tests par leur validité interne ou factorielle, qui indique dans quelle mesure les items constituant le test sont homogènes.» (Huteau; Lautrey, 1999, p.105).

2 Etudes métrologiques sur le Columbia

2.1. Travaux sur la sensibilité du Columbia dans les milieux américain et Français

En ce qui concerne la sensibilité du test, les auteurs de l'échelle qui ont travaillé, en 1954, d'abord sur une population américaine de 669 sujets de 5 à 23 ans dans un premier temps, puis, sur une deuxième population de 957 enfants, de 3 à 13 ans, ont souligné au fur et à mesure les qualités différenciatrices du test, qui permet de distinguer les enfants de différents âges, surtout dans l'intervalle de 3 ans et demi à 11 ans et demi. (Dague et coll., 1965)

Dans la population française les travaux de P. Dague et coll. (1965) sur un étalonnage de 400 enfants ont donné ces mêmes qualités différenciatrices entre les âges 3 ans et demi à 10 ans.

Les études sur la sensibilité, la fidélité et la validité de l'épreuve ont été reprises dans le manuel de 1972 de B. Burgemeister et coll. (Chaabane, 2008-2009) dans lequel ces questions de sensibilité, de fidélité et de validité sont discutées sur une population de 2600 enfants, de 3 ans et demi à 9 ans et 11 mois; une population représentative des Etats Unis pour laquelle les auteurs ont montré, encore une fois, la différence significative des valeurs entre les âges.

2.2 Travaux sur la fidélité du Columbia dans les milieux américain et syrien

La méthode half split a été calculée par les auteurs du Columbia à l'aide de l'équation de Spearman Brown. La corrélation s'est élevée entre 0.85 et 0.90. Ces mêmes auteurs ont noté une corrélation de 0.85 dans le test-retest.



W. Chabane (2008-2009) dans sa thèse de doctorat sur l'adaptation du Columbia à des populations aux besoins spécifiques en Syrie l'auteur a étudié les qualités métrologiques du test de Maturité Mentale sur une population de 790 enfants dont 118 enfants déficients.

En ce qui concerne la fidélité, l'auteur a utilisé la méthode split half et le test retest. Pour la split half, le kuder Richardson² donne chez les enfants déficients un coefficient de corrélation entre 0.58 et 0.69; et chez les enfants normaux un coefficient de 0.72 à 0.88. Le test retest a varié entre 0.65 à 0.92.

2.3 Travaux sur la validité du Columbia dans les milieux américain, européen et syrien

Selon P. Dague (1956) les études métrologiques sont peu nombreuses: Il cite Bessie Burgemeister et coll. (les auteurs du test en question) qui ont comparé le test du Columbia au STANFORD et à l'OTIS LENON.

La comparaison au STANFORD a donné une corrélation entre 0.44 et 0.60, Selon la nature des échelles employées. Quant à la comparaison au test OTIS LENON, elle a donné un coefficient de corrélation situé dans l'intervalle de 0.62 à 0.69

P. Dague a cité aussi Kodman, Waters et Whipple qui ont utilisé un échantillon de 207 enfants sourds et ont trouvé des corrélations de 0.61 à 0.73, avec les tests d'OTIS ALPHA et BETA. Quant à N. Chavannes et P. Dague (1965), ils ont trouvé une corrélation de Pearson de 0.83 avec le WISC chez 33 enfants myopathes.

La validité du Columbia, a été vérifié aussi par W. Chabane (2008-2009) dans son travail intitulé «Columbia Mental Maturity Scale; a study and standardisation of the scale on a sample of students with special needs» (thèse de doctorat en éducation sous la direction du Pr. Imtanius Mkhel, Damas).

Dans cette thèse sur l'adaptation du Columbia à des populations aux besoins spécifiques en Syrie, l'auteur a étudié

²-qui s'apparente au coefficient alpha de Cronbach.



les qualités métrologiques du test de Maturité Mentale sur une population de 790 enfants dont 118 enfants déficients.

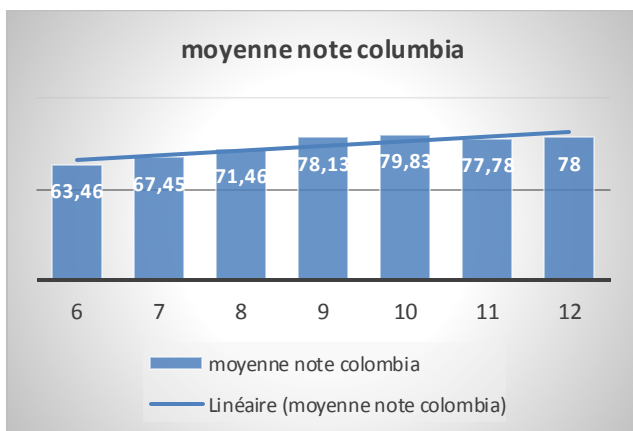
En ce qui concerne la validité du test Columbia, Chabane a utilisé le test d'OTIS (corrélation estimé à 0.35 chez les déficients et 0.58 chez les normaux, le WISC (corrélation estimé entre 0.53 et 0.78), la WIPSY (0.46 à 0.50), les progressive matrice de RAVEN (de 0.50 à 0.71). En conclusion, toutes les recherches précédentes démontrent la fiabilité et la constance du Columbia, chacune dans son propre milieu.

3. Etude de l'application du test Columbia en milieu algérien

3.1. La sensibilité du Columbia dans le milieu Algérien

Nous avons lancé une pré-enquête sur une population d'une centaine d'enfants afin de dégager les caractéristiques essentielles de l'utilisation et l'expérimentation des qualités métrologiques du test de maturité mentale en milieu algérien; nous parlerons ici principalement des étapes de l'expérimentation des concepts cités supra, à commencer par la sensibilité. Pour la population que nous avons examiné, 107 enfants de 6 à 12 ans, aussi bien les notes brutes que leurs transformation en quotient intellectuel les valeurs obtenues se différencient en fonction des âges des enfants, tel qu'on peut le voir sur le graphe.

Figure 1: Sensibilité du Columbia dans le milieu algérien



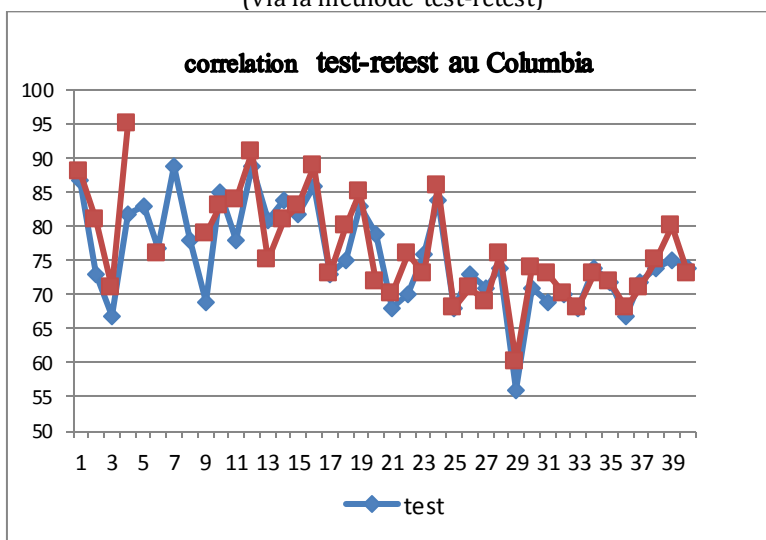
On voit bien sur le graphe que les différences sont très bien délimitées entre les âges de 6 à 10 ans, mais elles ne sont plus significatives entre les années à partir de l'âge de 10ans. La même situation a été signalée dans l'étalonnage français puis a été revue par les auteurs du test, qui ont fini par arrêter leur nouvel étalonnage à l'âge de 9ans et demi. (Manuel de 1972)

3.2 La fidélité du Columbia dans le milieu algérien

La question de la fidélité du Columbia a été vérifiée dans notre recherche par l'utilisation de 03 techniques: le test-retest, la halph split et l'alpha de Cronbach.

Pour le test -retest, nous l'avons vérifié sur une population de 60 enfants, algériens, de 6 à 11ans, et nous avons obtenu un coefficient de corrélation de Pearson de 0. 85 (voir graphe):

Figure 2: fidélité du Columbia dans le milieu algérien
(via la méthode test-retest)



Il apparait clairement du graphe ci-dessus que le test a fourni pratiquement le même résultat chez le même sujet au cours des deux passations du Columbia. Et avec un coefficient de corrélation (de Pearson) de .0 85, on ne peut que dire que la cohérence interne est très bonne.

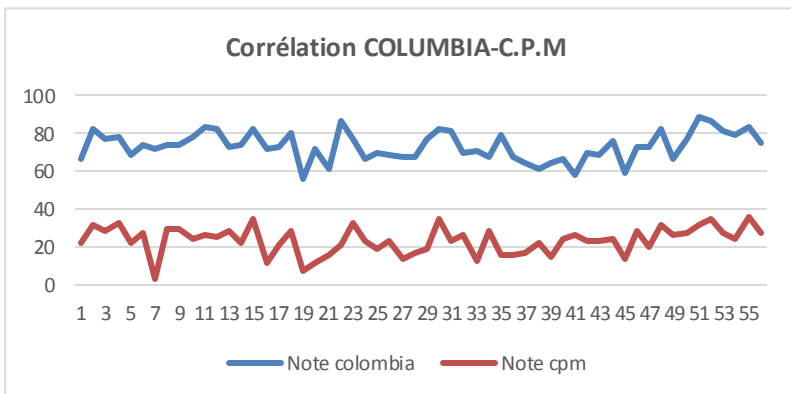


- Pour la half Split, elle a été calculée sur une population de 500 enfants algériens, de 3 à 11ans, et sa valeur est de 0.717, ce qui indique une bonne cohérence interne.
- L'alpha de Cronbach aussi a été calculé sur la même population (500 enfants algériens, de 3 à 11ans), et la valeur de Alpha est de 0.936, ce qui indique une excellente cohérence interne.

3.3 La validité du Columbia dans le milieu Algérien

Nous avons à notre tour évaluer la validité du Columbia dans le milieu algérien, par la comparaison des valeurs du test du Columbia avec les Colored Progressive Matrice de RAVEN (CPM) sur une population de 56 enfants, tel qu'il apparait dans le graphe suivant:

Figure 3: validité du Columbia dans le milieu Algérien
Corrélation Columbia-CPM



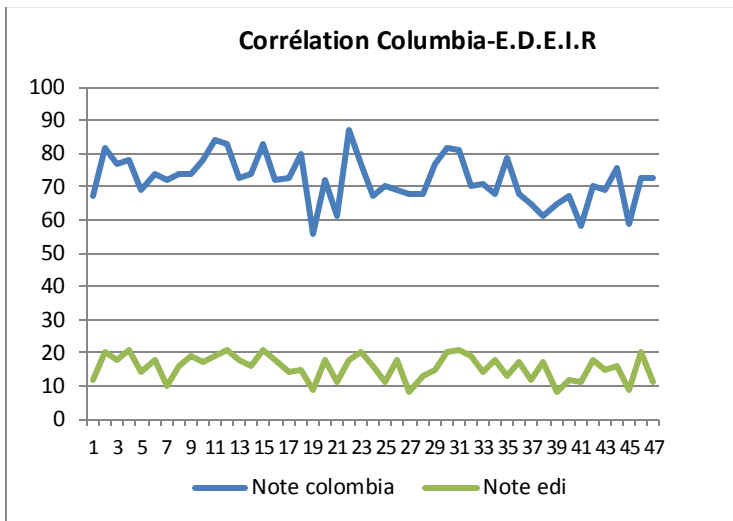
Le coefficient de corrélation a été estimé à 0 .57, un coefficient assez appréciable, sachant que les CPM renvoient plus au raisonnement abstrait alors que le Columbia renvoie à l'analyse catégorielle et classificatoire.

La validité, dans notre recherche a été également calculée par comparaison des valeurs du test du Columbia avec l'Echelle Classification, des Echelles d'Efficiency Intellectuelles Révisée de M. Perron Borelli (2000);qui est décrite par son auteur comme en rapport avec l'analyse catégorielle et classificatoire.



Dans ce cadre le coefficient de corrélation de Pearson est de 0.66, ce qui est tout à fait respectable en ce qui concerne sa fiabilité.

Figure 4: validité du Columbia dans le milieu Algérien
Corrélation Columbia-E.D.I.R.



La comparaison du test du Columbia avec l'échelle classification, des échelles d'efficience intellectuelles révisée semble plus appropriée, et confirme d'avantage la validité du Columbia, tel qu'il apparait dans le graphe, probablement parce que les deux sont des tests en rapport avec l'activité classificatoire.

Conclusion

Le travail d'expérimentation que nous avons mené sur le terrain nous a permis d'atteindre plusieurs objectifs notamment la démonstration des qualités différenciatrices du test de maturité mentale du Columbia.

De ce fait, et en plus des recherches expérimentales menées ailleurs, on peut affirmer que jusqu'à l'âge de 10 ans, le Columbia est un outil de différenciation, qui délimite parfaitement entre les âges, que ce soit dans le milieu algérien ou ailleurs.



La fidélité du test de maturité mentale du Columbia en milieu algérien n'est plus discutable non plus, puisqu'elle a été vérifiée dans notre recherche par l'utilisation de 03 techniques: le test – retest, la halph split et l'alpha de Cronbach. Toutes ces techniques ont démontré que le Columbia est un test qui se distingue par sa stabilité et sa constance.

On a aussi pu vérifier la validité du Columbia, c'est-à-dire qu'il renvoie parfaitement à l'analyse catégorielle et classificatoire, ceci par la comparaison des valeurs du test du Columbia avec d'autres tests, particulièrement avec l'Echelle Classification, des Echelles d'Efficiences Intellectuelles Révisée de M. Perron Borelli (2000), qui est lui aussi en rapport avec l'activité classificatoire. Ceci rend le test de maturité mentale du Columbia parfaitement praticable dans la société algérienne.

Bibliographie

1. Allaire D., Robert M., 1988. *Fondements et étapes de la recherche scientifique en psychologie*, Québec Canada, Edisem.
2. Chabane W., 2008-2009. *Columbia mental maturity scale; a study and standardisation of the scale on a sample of students with special needs*, thèse de doctorat en éducation sous la direction du Pr Imtanius Mkhel, Damas.
3. Dague P., et coll. 1965. *Echelle de maturité mentale de Columbia*, CREAPSY, Alger 2007.
4. Ghillevic C., Vautier S., 1998. *Diagnostic et tests psychologiques*; Paris, Nathan.
5. Huteau M.; Lautrey J., 1999. *Evaluer l'intelligence; psychométrie cognitive*, Paris PUF.
6. Pichot P., 1999. *Les tests mentaux*, PUF.
7. Reuchlin M., 1962. *Les méthodes quantitatives en psychologie*, Paris, PUF.
8. Reuchlin M., 1964. *Méthodes d'analyse factorielle à l'usage des psychologues*, Paris, PUF.
9. Reuchlin M., 1969. *La psychologie différentielle*, Paris, PUF
10. Reuchlin M., 1969. *Les méthodes en psychologie*, Paris, PUF.

