

تقنين الاختبارات في البحوث النفسية والتربوية

بوقرة عواطف *

جامعة المسيلة. الجزائر

Awatif.bouguerra@univ-msila.dz

تاريخ الارسال : 2022/04/01

تاريخ القبول: 2022/04/15

الملخص:

سعت الدراسة الحالية إلى التعريف بعملية تقنين الاختبارات في البحوث النفسية والتربوية؛ باعتبارها أهم خطوة من خطوات القياس النفسي والتربوي في البحث العلمي. نظرا لكونها تؤكد صلاحية أداة القياس المستخدمة وتتحقق من صدقها وثباتها.

كما سعت إلى توضيح خطوات عملية التقنين. وتوصلت إلى أنه هناك بعض الاختلافات في هذه العملية بين البحوث النفسية والتربوية إذ أنه لا يتم الاعتماد على قياس صعوبة وسهولة الفقرات في بعض الاختبارات النفسية إلا ما تعلق منها بدراسة القدرات العقلية. كما أنه لا يتم دراسة فعالية الموهبات ومؤشراتها في الاختبارات النفسية عكس الاختبارات التربوية التي تعتمد على قياس التحصيل الدراسي في أغلب الدراسات البحثية.

كما توصلت إلى تأكيد أهمية خطوة التحقق من مستوى خطأ القياس في الدراسات النفسية والتربوية على حد سواء لما له من أهمية.

الكلمات المفتاحية: تقنين الاختبارات، البحوث النفسية، البحوث التربوية.

مقدمة:

تختلف أدوات القياس باختلاف الظاهرة المدروسة سواء أكانت تربوية أو نفسية. فقد نعتمد المقابلة والملاحظة، الاختبارات ؛ هذه الأخيرة التي تعنى بقياس الظاهرة المدروسة في غالب الأحيان دراسة كمية؛ تعطينا درجة وجود وتوفر السمة المقاسة لدى المفحوص.

* المؤلف المرسل: بوقرة عواطف، الايميل: awatif.bouguerra@univ-msila.dz

تحتوي مجموعة من البنود تمثل لنا مؤشرات سلوكية يمكن ملاحظتها وقياسها إجرائيا في الميدان، ويمكن للمفحوص تحديد درجة توافرها لديه.

ونظرا لاعتمادها في عملية الفحص والتشخيص وجب التحقق من صلاحيتها من خلال اخضاع بنودها للتحليل الاحصائي والتحقق من صدقها وثباتها وموضوعيتها في القياس. وهذا ما يعرف بعملية التقنين؛ هذه الأخيرة التي تعد أهم خطوة من خطوات بناء وتصميم الاختبارات سواء النفسية أو التربوية في البحث العلمي.

لشرح هذه العملية بالتفصيل سنعمل من خلال هذا البحث على التعريف بها، وذكر أهدافها، وخطواتها. والتوصل في النهاية إلى ذكر بعض الاختلافات الواردة في عملية التقنين بين الاختبارات النفسية و الاختبارات التربوية.

أولا: التعريف بعملية التقنين والاختبار المقنن

التعريف بعملية التقنين:

تعددت تعريف الباحثين لعملية تقنين الاختبار سواء النفسي أو التربوي إلا أنها تتفق في بعض النقاط . نذكر من بين هذه التعاريف ما يلي:

تعريف زيدان (1979) هو : " العملية التي يتم من خلالها التحكم في العوامل غير المناسبة التي يمكن أن تؤثر في عملية القياس؛ وذلك من خلال تخفيض أخطاء القياس إلى حدها الأدنى عن طريق اختيار عينة ممثلة لمجتمع الدراسة يطبق عليها الاختبار؛ ومن ثم توحيد فقراته وإجراءات تطبيقه وتصحيحه بشكل يوفر للاختبار خصائص سيكومترية تتفق مع خصائص الاختبار الجيد، ومن ثم توفير المعايير المناسبة لتفسير الدرجات الخام". (زيدان، 1979، ص 68)

تعريف طه (2007): : " هي مختلف الدراسات النظرية والميدانية والتجريبية والإحصائية التي تجرى على الاختبار أو المقياس النفسي حتى نطمئن من صلاحيته للاستخدام؛ وبذلك نثق في نتائجه مثل دراسة وحداته وثباته وصدقته وتحديد معايير ووضع نماذج تصحيحه وكيفية تفسيره وتحديد كيفية تطبيقه وتعليماته". (طه، 2007، ص 258.259)

تتفق التعاريف السابقة على أن عملية التقنين هي:

- توحيد فقرات الاختبار النفسي والتربوي وإجراءات تطبيقه وتصحيحه بشكل يوفر للاختبار خصائص سيكومترية تتفق مع خصائص الاختبار الجيد من موضوعية وصدق وثبات. بهدف التخفيف من أخطاء القياس.

2 تعريف الاختبار المقنن:

عرفه Borg and gall بأنه "ذلك الاختبار الذي يحافظ على صدقه (أي قياس ما أعد لقياسه) وثباته (أي الوصول إلى النتائج نفسها لو تكرر تطبيقه) خاصة إذا اتبعت التعليمات المصاحبة له. " (العساف، 2000، ص 5)

و عرفه بوسنة محمود (2007): "بأنه أداة وصفية تكيفية لظاهرة سلوكية معينة ميزتها الأساسية أنها مقننة". (بوسنة محمود، 2007، ص 771)

تعريف سعد عبد الرحمن (1979): هو مجموعة من الأسئلة أو البنود التي تدور حول موضوع واحد أو عدة مواضيع وليس لها إجابات صحيحة أو خاطئة إذ أن المطلوب فيها معرفة رأي الفرد أو نوعية استجابته في موقف من المواقف التي يمثلها ذلك السؤال أو البند. (سعد عبد الرحمن، 1998، ص 760)

يتضح من خلال التعاريف السابقة أن الاختبار المقنن:

- عبارة عن مجموعة من الأسئلة تسعى إلى التعرف على آراء الأفراد في مختلف المواقف.
- هو الاختبار الذي تم التحقق من صدقه وثباته بهدف قياس ظاهرة سلوكية معينة (تربوية، نفسية، اجتماعية ..)

ثانيا: اهداف عملية التقنين :

تهدف عملية تقنين الاختبار إلى:

- توحيد جميع إجراءات الاختبار و المتمثلة في: إجراءات التطبيق ، التصحيح ، و تفسير الدرجات (اسماعيل .2004: 31)
- حساب معاملات صدق الاختبار
- حساب معاملات ثبات الاختبار

- استخراج معايير موحدة للاختبار (المتوسطات الحسابية- الانحرافات المعيارية- الدرجات المعيارية) (معمرية. 2012: 296)

ثالثا: الخطوات الأساسية في عملية التقنين

تعد عملية التقنين الخطوة الأخيرة من مراحل تصميم الاختبار النفسي والتربوي واعداده. حيث وجب: أولاً: ضبط اجراءات اعداد الاختبار وصياغة بنوده، وطريقة تقديمه وتصحيحه وطريقة تفسيره موحدة في كل المواقف ؛ وبالتالي يمكن الحصول على نفس النتائج رغم تعدد الفاحصين، اضافة إلى تحديد الخصائص السيكمترية التي تدل على جودة الاختبار. (مراد وسلمان، 2005، ص 349)

ثانياً: تقنين الاختبار على عينة ممثلة للمجتمع تعرف بعينة التقنين ؛ للحصول على معايير تسمح بتفسير الدرجات الخام التي يحصل عليها المفحوصين.

ويلخص (الدوسري، 1999) خطوات عملية بناء وتقنين الاختبار كما يلي:

- 1- تحديد المجتمع الذي سيقنن عليه الاختبار تحديداً إجرائياً دقيقاً: وهي الخطوة الأساس لضمان صحة الخطوات اللاحقة في عملية التقنين، وتتضمن تحديد أهم خصائص وسمات المجتمع الديموغرافية من حيث الخصائص الجغرافية والسكانية والاقتصادية والتعليمية.....
- 2- اختيار العينة الممثلة للمجتمع وتحديد أسلوب اختيارها: وتعتمد هذه الخطوة على سابقتها؛ لأن اختيار حجم العينة وأسلوب اختيارها يعتمد بشكل كبير على المعلومات المتوفرة عن مجتمع الدراسة، وبصفة عامة فكلما كان حجم العينة كبيراً كلما كان أفضل وأقرب إلى التمثيل الجيد للمجتمع. ويختلف الأسلوب الذي نختار به عينة التقنين من المجتمع المستهدف؛ حيث يبدأ من طريقة سحب عينة عشوائية بسيطة حتى يصل إلى استراتيجيات سحب العينات الأكثر تعقيداً مثل العينة العشوائية الطبقية أو العينة العنقودية .

3- التخطيط الجيد والمسبق لتطبيق الاختبار: وذلك بوضع خطة شاملة لتطبيق الاختبار

تتضمن تحديد الإجراءات والخطوات التي سوف تتبع، وتجهيز جميع أدوات ومستلزمات الاختبار، مع وضع قوائم بأسماء الأماكن التي سوف يتم تطبيق الاختبار فيها ؛ مع وضع برنامج زمني للتنفيذ.

- 4- تطبيق الاختبار: وتستلزم هذه الخطوات توحيد ظروف إجراء تطبيق الاختبار لجميع أفراد العينة، وذلك لتحقيق مبدأ تكافؤ الفرص أمام الجميع لضمان أن الفروق التي رصدها الاختبار تعود للفروق في أداء الأفراد فقط .
- 5- تحليل فقرات الاختبار: وذلك للتعرف على مدى فعالية فقرات الاختبار، ومدى إسهامها في الحصول على خصائص سيكومترية (صدق وثبات) تتفق مع خصائص الاختبار الجيد وتشمل هذه الخطوة حسب (أبو ناهية، 1994) التعرف على ما يلي:
- 5-1 معامل صعوبة الفقرات **Difficulty Coefficient**: يشير معامل صعوبة الفقرة إلى النسبة المئوية لعدد المفحوصين الذين أجابوا بشكل صحيح على الفقرة، وهو يقع بين الصفر والمائة .
- 5-2 معامل سهولة الفقرات **Ease Coefficient**: يشير معامل سهولة الفقرة إلى النسبة المئوية لعدد المفحوصين الذين أجابوا بشكل خاطئ على الفقرة.
- 5-3 القدرة التمييزية لكل فقرة من الفقرات: أشار "معامل تمييز الفقرة إلى قدرة الفقرة على التمييز بين المجموعات المتباينة أي بين المجموعتين العليا والدنيا. ويتم حساب معامل التمييز باستخدام معامل الارتباط المنصف للسلسلة أو بنسبة 27 % من المجتمع الكلي.
- 5-4 فعالية المشتتات (المموهات) لكل فقرة من الفقرات: يقصد بفعالية المشتتات أو المموهات هي قدرة هذه المشتتات على جذب المفحوصين إليها. والمموه الجيد أو الفعال يجب أن يجذب إليه عدداً أكبر من المفحوصين من بين أفراد المجموعة الدنيا وعدداً أقل من المفحوصين من بين أفراد المجموعة العليا. بينما البديل أو الاختيار الصحيح يجب أن يجذب إليه معظم المفحوصين الذين يعرفون الإجابة الصحيحة من بين أفراد المجموعتين العليا والدنيا.
- 6- التحقق من ثبات الاختبار: **Reliability of test** الذي يشير إلى استقرار الدرجات التي يحصل عليها الفرد في اختبار ما. إذا ما أعيد تطبيقه بعد فترة زمنية معينة تحت نفس الظروف. (Groth Marnat,2009 ,p12)

ويشير إلى درجة الدقة أو الضبط والإحكام في عملية القياس بحيث تعطينا معاملات الثبات فكرة عن درجة الاتساق أو التوافق في نتيجة القياس عند تكراره. ويقصد بثبات الاختبار دقة الاختبار في القياس أو الملاحظة وعدم تناقضه مع نفسه. فالاختبار الثابت هو الذي يعطي نفس النتائج (تقريباً) إذا طبق على نفس الأشخاص في موقعين مختلفين. ويعبر عنه إحصائياً بحساب معامل الارتباط بين الدرجات التي حصل عليها المفحوص في المرة الأولى وبين النتائج التي حصل عليها في المرة الثانية.

6-1 طرق حساب معامل الثبات: اختلفت طرق حساب معامل الثبات للقياس وذلك بناء على نوع القياس والغرض منه ومدى الخطأ المعياري لكل قياس، وقد ظهرت عدة طرق وأساليب إحصائية لحساب معامل الثبات نستعرض منها الطرق التالية :

طريقة تطبيق و إعادة تطبيق الاختبار نفسه: Test – Retest Reliability تقوم فكرة هذه الطريقة على حساب معامل الاستقرار (الثبات) من خلال تطبيق الاختبار على مجموعة من الأفراد تتراوح بين (10 إلى 40 فرد) ثم إعادة التطبيق على نفس الأفراد وتحت نفس الظروف بعد فاصل زمني يتراوح بين أسبوع وشهر. ويتحدد الفاصل الزمني بين التطبيقين وفق نوع التفسير المطلوب للدرجات. ويكون معامل الثبات هو معامل الارتباط البسيط بين درجات الاختبار في التطبيقين الأول والثاني. (السيد حسين، 2006، ص 4)

هناك مجموعة من الصعوبات التي تكتنف استعمال هذه الطريقة ويمكن ذكرها كما يلي:

- الفترة الفاصلة بين التطبيق الأول و الثاني: لو طبقنا اختبار معين على فرد ما، ثم أعدنا تطبيقه بعد فترة زمنية معينة، فإن الخبرة التي أكتسبها الفرد في المرة الأولى سوف تؤثر على أدائه في المرة الثانية
- طبيعة السمة المقاسة: فرضاً لو طبقنا اختبار ثم أعدنا تطبيقه بعد فترة زمنية فإنه قد يحدث تغير في السمة وضع القياس مما يؤدي بنا إلى الظن أن الأداة التي نستخدمها غير ثابتة، لذا عند استعمال طريقة التطبيق وإعادة التطبيق يجب أن لا تكون الفترة بين التطبيقين بالطويلة أو القصيرة ويعاب على هذه الطريقة خضوعها لكل ما يؤثر في المختبرين من خلال الفترة الزمنية الفاصلة بين التطبيقين.
- صعوبة ضبط الظروف في التطبيقين .

طريقة الصور المتكافئة: Equivalent – Form Reliability يتم تصميم صورتين متكافئتين من الاختبار الواحد والتكافؤ كما حددته الدكتور "منال مبارز" يشمل على الجوانب التالية:

أن تكون المجالات والموضوعات التي تقيسها الصورتان واحدة.

تساوى نسبة المفردات التي تخص كل موضوع من الموضوعات.

تماثل مستوى صعوبة الفقرات.

تشابه طريقة صياغة الفقرات.

تساوى طول الصورتين وطريقة إجرائهما وتقدير درجتهما وتوقيتهما.

تساوى متوسط وتباين درجات الأفراد في الصور المتكافئة .

تتميز هذه الطريقة بعدم التأثير بالتدريب أو الخبرة من صورة لأخرى، ويختفي أثر الألفة لاختلاف بنود الصورتين، كما أنها تصلح لحساب معامل ثبات اختبارات الذاكرة والاختبارات التحصيلية.

ومن عيوب هذه الطريقة صعوبة تصميم اختبارين متكافئين لقياس نفس السمة أو الخاصية. وانتقال أثر التدريب وبخاصة في حالة اقتران عملية تطبيق الصورتين، وقصر الفترة الزمنية أو طولها بين تطبيق الصورتين المتكافئتين، والجهد المضاعف المبذول والوقت والتكلفة المبذولة في بناء الاختبار المكافئ. (السيد حسين، 2006، ص.ص، 5.4)

طريقة التجزئة النصفية: Spilt–Half Reliability

تعتمد هذه الطريقة على تجزئة الاختبار المطلوب تعيين معامل ثباته إلى نصفين متكافئين وذلك بعد تطبيقه على مجموعة واحدة. وفي هذه الطريقة يطبق الاختبار مرة واحدة فقط ثم تقسم درجات العينة إلى نصفين متكافئين تماماً من حيث العدد، ومستوى السهولة، والصعوبة. ولكي يتحقق ذلك فإنه ينبغي أن يقسم الاختبار ، وتوجد عدة طرق لتقسيم الاختبار حسبما أورده (الزويني، 2014) :

القسم النصفية: وذلك بقسمة الاختبار إلى نصفين متساويين فمثلاً لو كان لدينا اختبار يتكون من (60) فقرة فسيكون النصف الأول (1-30) فقرة والثاني من (31-60) فقرة

التقسيم الفردي والزوجي: يتم ذلك بقسمة فقرات الاختبار إلى نصفين بحسب أرقام الفقرات فالفقرات ذات الأرقام الفردية مثلا تعد قسما والفقرات ذات الأرقام الزوجية قسما آخر .

التقسيم الجزئي: تشمل بعض الاختبارات على اختبارات فرعية وفي مثل هذه الحالة لا يمكن استخدام أي نوع من الأنواع أعلاه لذلك يقسم الاختبار إلى جزئين أو ثلاثة أجزاء بحسب ما يحتويه الاختبار من جوانب فرعية فمثلا إذا كان الاختبار من (6) اختبارات فرعية فيكون القسم الأول من الفروع (1-3-5) والقسم الثاني من الفروع (2-4-6) وهكذا ان معامل الارتباط المستخرج بين نصفي الاختبار يفسر بأنه معامل الاتساق الداخلي. ولما كان معامل الثبات لا يقيس التجانس الكلي للاختبار لأنه ثبات لنصف الاختبار فهناك طرق متعددة لتلافي هذا النقص وذلك باستخدام معامل سبيرمان براون ومعادلة رولون وغيرها. (بوقرة، 2021، ص 59)

7- التحقق من صدق الاختبار: Test Validity

يدل الصدق على قدرة المقياس على قياس ما أعد لقياسه. أي قدرته على قياس ما يدعي قياسه من جوانب سلوك الأفراد.

أو هو درجة قياس الاختبار لما وضع لقياسه.

ولكي يكون الاختبار صادقا لا بد من توافر شرطين أساسيين هما :

- أن يناسب الاختبار طبيعة العينة المراد دراستها.

- أن يقيس الاختبار الصفة أو الخاصية التي وضع لقياسها. (بوقرة، 2021، ص 71)

7-1 طرق حساب صدق الاختبار:

سنعمل أولا على ذكر مجموعة من طرق القياس الكلاسيكية الأكثر رواجاً في الدراسات والبحوث الأكاديمية النفسية والتربوية. وسنوضح لاحقا الأنواع التي تركز عليها النظرية الحديثة في القياس.

تم تقسيم طرق التحقق من صدق الاختبار وفق النظرية الكلاسيكية للقياس إلى قسمين قسم لا يعتمد على الاختبارات المحكية المرجع، وقسم يعتمد على الاختبارات محكية المرجع.

أولاً: أنواع الصدق وفق النظرية الكلاسيكية

1-1-1 طرق القياس التي لا تعتمد على اختبارات محكية المرجع :

صدق المحكمين: أو الصدق الظاهري. يمكن حساب الصدق الظاهري للاختبار عن طريق التحليل المبدئي لفقراته من خلال عرضه على مجموعة من المحكمين والأخصائيين لتحديد ما إذا كانت هذه الفقرات تقيس وتنتمي للمجال الذي وضعت لقياسه.

ثم يقوم الباحث بحساب تكرارات استجابات المحكمين وإصدار الحكم بناء على النتائج التالية:

قبول العبارة: في حال كانت نسبة الاتفاق أكبر من أو تساوي 60%.

حذف العبارة: في حال كانت نسبة الاتفاق أقل من 40%.

تعديل العبارة: في حال كانت نسبة الاتفاق تتراوح بين 40 إلى 60%.

قبول المقياس ككل: إذا كانت نسبة الاتفاق أكبر من أو تساوي 77%.

لا تعد هذه الطريقة كافية للتأكد من صدق الاختبار. ولكنها تفيد في طمأنة الباحث مبدئياً على دقة الاختبار الذي يستخدمه في مقياس ما وضع لقياسه .

صدق المقارنة الطرفية: في هذه الطريقة نعمل على تقسيم المقياس إلى قسمين: قسم في الطرف الأعلى يضم نسبة 27 % أو 33 % من البنود. وقسم في الطرف الأدنى يضم نسبة 27 % أو

33 % من البنود. ثم نقوم بحساب قيمة " ت " من خلال حساب اختبار " T- Test " للفروق بين عينتين مستقلتين متساويتين. باعتماد المعادلة التالية :

والتحقق من دلالة قيمة " ت " من خلال مقارنة القيمة المحسوبة مع القيمة الجدولية.

في حال كانت هناك فروق بين القسم الأعلى والقسم الأدنى يمكننا القول بأن المقياس صادق وقابل للتطبيق على مجتمع الدراسة وبالتالي حقق شروط الصدق الخارجي .

الصدق التمييزي: يعمل على تقسيم أفراد العينة إلى قسمين قسم أعلى وقسم أدنى بأخذ عينة 27 % أو 33 % . وحساب معاملات التمييز لفقرات المقياس بناء على إجابات القسمين.

بتطبيق المعادلة التالية:

معامل التمييز = عدد الإجابات الصحيحة عن الفقرة في المجموعة العليا _ عدد الإجابات الصحيحة عن الفقرة في المجموعة الدنيا ÷ عدد أفراد أحد المجموعتين.

بحيث يتم قبول الفقرات بناء على مؤشرات التمييز التالية:

- إذا كان معامل التمييز أكبر من 0.40 فإن الفقرة تعتبر ذات تمييز عالي وممتاز.
- إذا كان معامل التمييز بين (0.30 - 0.39) فإن الفقرة تعتبر ذات تمييز جيد.
- إذا كان معامل التمييز بين (0.20 - 0.29) فإن الفقرة تعتبر ذات تمييز جيد إلى حد ما (تحتاج إلى تعديل).

- إذا كان معامل التمييز أقل من 0.19 فإن الفقرة ضعيفة وينصح بحذفها.

7-1-2 طرق القياس التي تعتمد على اختبارات محكية المرجع: يعتمد في أغلب الأحيان على حساب معاملات الارتباط كمؤشرات للثبات الداخلي، ودلالاتها كمؤشرات للثبات الخارجي.

الصدق التنبؤي :

هو مدى قدرة الاختبار على التنبؤ بنتيجة معينة ويستخدم هذا النوع من الصدق في اختبارات الاستعداد التي تهدف في الأصل إلى التنبؤ بما يمكن أن ينجزه المفحوص في المجال الدراسي أو المهني في المستقبل.....

فالصدق التنبؤي يعني مدى صلاحية الاختبار في الإشارة إلى مستوى الانجاز الذي ينتظر أن يصل إليه المفحوص في موضوع مماثل أو قريب من حيث نوعه أو صعوبته والوصول إلى هذا الصدق يساعد في التحقق من قدرة الاختبار على تحقيق أحد أغراضه التي وضع من أجلها وهي اتخاذ القرارات حول المستقبل.

يجب مراعاة ثلاث أبعاد في هذا النوع من الصدق :

- حساب القيمة التنبؤية للاختبار .
- الاعتماد على فكرة أن السلوك له صفة الثبات النسبي في المواقف المستقبلية .
- التنبؤ يحتاج إلى فترة بين تطبيق الاختبار ثم جمع البيانات عن المحك في فترة تالية للاختبار (الطريقة التتبعية).

فالصدق التنبؤي يعني عبارة عن إمكانية التنبؤ بنتائج لاحقة والتي افترضها القائم بالاختبار مسبقاً

الصدق التلازمي: يعد من أنواع الصدق التجريبي. ويستخدم عندما يتلزم تطبيق الاختبار وتطبيق المحك معاً ويصبح الهدف هو معرفة ما إذا كان كل من الاختبارين يقيسان خصائص قائمة بالفعل في وقت واحد، وذلك بهدف تقدير الحالة الراهنة، وهو من أنسب الأساليب ملائمة للاختبارات التشخيصية .

يتطلب هذا النوع من الصدق التأكد من أن الاختبار النفسي يسمح بإبراز ارتباط مختلف المتغيرات كما تنبأت به النظرية التي قام الاختبار على أساسها أي ما إذا كان الاختبار حساساً حساسية كافية لقياس آثار سَلَمَت بها النظرية التي ساعدت على بنائه وتصميمه. يكون الاختبار صادقاً تلازمياً لما تكون الارتباطات كبيرة .

هناك عدة طرق نستخدمها في الصدق التلازمي :

-**الصدق التقاربي:** يعتمد على إعطاء اختبار يقيس نفس الظاهرة التي يقيسها الاختبار الأول : فمثلاً عندما يصمم باحث اختبار لقياس اتجاهات الطلبة نحو عمل المرأة ويريد ان يتعرف على صدقه، فانه يقوم بإعطاء اختبار آخر يتناول نفس الظاهرة إلى عينة الطلبة نفسها ثم يستخرج العلاقة الارتباطية بين درجات الطلبة في الاختبارين ويكون الاختبار صادقاً تلازمياً عندما تكون الدرجة الارتباطية عالية بين الاختبارين .

-**الصدق التباعدي:** يعتمد على المجموعات المضادة: إن الاختبار الصادق هو الذي يميز بين الأفراد فاختبار الذكاء يجب أن يميز بين الأذكياء والأقل ذكاء أو اختبار القلق يجب أن يميز بين الفرد القلق وغير القلق وعلى هذا الأساس يختار مصمم الاختبار مجموعتين من الأفراد تختلفان في الظاهرة التي يقيسها الاختبار، ويعطي لهما الاختبار فإذا كانت هناك فروق واضحة بين إجابتي المجموعتين في الاختبار، فإنه يكون قد تحقق فيه الصدق التلازمي .

أو نقوم بقياس سمة معاكسة للسمة المراد قياسها من الاختبار. مثلاً قياس مستوى التفاؤل باختبار وقياس مستوى التشاؤم باختبار آخر. ثم نقوم بحساب معامل الارتباط بين الاختبارين فإذا كانت العلاقة عكسية دل ذلك على أن الاختبار المراد التحقق من صدقه صادق.

ثانياً: أنواع الصدق وفق ما تراه النظرية الحديثة للقياس

طرق التحقق من الصدق البنائي: (صدق التكوين الفرضي، صدق المفهوم)

نعمد للتحقق من الصدق البنائي للاختبار على استخدام طرق الارتباط، الفروق، التجريب كما هو موضح فيما يلي:

الطرق التي نعمد على الارتباط:

الاتساق الداخلي: نعمد على حساب معامل الارتباط بين السؤال والبعد الذي ينتمي إليه. أو بين الأبعاد الفرعية المكونة للأداة والدرجة الكلية للأداة. أو حساب الارتباط بين السؤال والدرجة الكلية للاختبار.

مصنوفة السمات: أو الطرق المتعددة نعمد حسب Campbell And Fiske 1959 على :

الصدق التقاربي: إذا كان الارتباط عالي بين أداة القياس وأدوات أخرى تقيس نفس السمة.

الصدق التباعدي (التمايزي) : إذا كان الارتباط ضعيف بين المقياس ومقاييس أخرى لا تقيس نفس السمة.

ثانياً: الطرق التي نعمد على الفروق

المقارنة الطرفية :

الطريقة الأولى: مقارنة الأطراف في المقياس والمحك الخارجي.

الطرف الأعلى للمقياس مع الطرف الأعلى للمحك.

والطرف الأدنى للمقياس مع الطرف الأدنى للمحك.

يكون المقياس صادق إذا لم يكن هناك فروق ذات دلالة إحصائية بين متوسطات الدرجات العليا والدنيا للمقياس.

الطريقة الثانية: مقارنة الأطراف في المقياس فقط بالاعتماد على درجات الثلث الأعلى والثلث الأدنى من المقياس (27% أعلى 27% أدنى) وحساب الفرق بين المتوسطين.

إذا كانت هناك دلالة إحصائية بين المتوسطين دل ذلك على صدق المقياس. (سعد عبد الرحمن، 191، 1998)

الفروق بين المجموعات: من خلال دراسة الفروق بين مجموعات متضادة (مختلفة).

ثالثا: الطرق التي تعتمد على التجريب

يتم من خلال اختبار الفرض القائل بتغير درجات الاختبار بتغير أوضاع المعالجات التجريبية. حيث تتم أكثر من معالجة تجريبية في مواقف مختلفة. فإذا تغيرت الدرجات والنتائج وفقا لذلك (التجريب) دل على أن الاختبار يقيس السمة أو الخاصية المراد قياسها. (علام، 2000، ص 219)

8- تحديد الخطأ المعياري للمقياس:

يشير (عودة، 2004) أن القياس لا يمكن أن يكون مثاليا إلا بتقدير العلامة الحقيقية للقياس باعتماد العلامة الظاهرية. أين يمكن أن تقع العلامة الحقيقية ضمن مدى معين للعلامة الظاهرية ويشار إلى هذا المدى بفترة الثقة. ولتحديد هذه الفترة فإننا بحاجة إلى معرفة الخطأ المعياري في القياس ويمكن تقدير الخطأ المعياري من خلال معامل الثبات والانحراف المعياري للعلامات الظاهرية.

9- اشتقاق المعايير للمقياس:

تدعى النتائج التي نحصل عليها من تطبيق الاختبارات والقياسات درجة خام، وتعرف بأنها النتيجة الأصلية المشتقة من تطبيق الاختبارات أو أي أداة قياس أخرى قبل أن تعالج إحصائيا . وعليه لابد من التعامل إحصائيا مع الدرجة الخام لتحويلها إلى درجة معيارية؛ والدرجة المعيارية حسب (الداودي، 2013) هي درجة يعبر فيها عن درجة كل فرد على أساس عدد وحدات الانحراف المعياري لدرجته عن المتوسط. ويطلق عليها أحيانا اسم (المسطرة). (بوقرة، 2021، ص 83)

ومن أشهر الدرجات المعيارية (الدرجة الزائفة ، النائية ، المثبينة) . ووجب اشتقاق معايير المقياس ووضعها في شكل جداول في دليل الاختبار للتمكن من الحكم على نتائج المفحوصين المتوصل إليها بعد تطبيق الاختبار لتشخيص حالاتهم.

رابعاً: الاختلافات الواردة في عملية تقنين الاختبارات النفسية والتربوية

تشترك الاختبارات النفسية والتربوية في عملية التقنين. إلا أنها تختلف في بعض الأمور التفصيلية الناتجة عن طبيعة السمة التي تفرض علينا قياسات معينة ولا تفرض علينا قياسات أخرى. وفيما يلي توضيح لبعض هذه الأمور التفصيلية:

- عدم وجود قياس معاملات الصعوبة والسهولة للاختبارات النفسية إلا في حالة القياسات العقلية. وذلك لكون البنود لا تحتل وجود اجابات صحيحة وأخرى خاطئة. لأننا نتكلم عن سمات شخصية وأمزجة لا يوجد فيها الصح والخطأ.
- عدم وجود مؤشرات للمموهات في الاختبارات النفسية: لأنه لا توجد مموهات ضعيفة ، فاعلة، وخاطئة. وذلك لعدم وجود مموهات صحيحة في الاختبارات النفسية.
- استخراج القوة التمييزية للفقرات سواء كان الاختبار نفسي أو تربوي، وذلك للاعتماد عليها في عملية الفحص والتشخيص.
- اختلاف عينة التقنين: فالاختبارات التربوية تشمل عينات كبيرة نتيجة لطبيعة السمة المدروسة ،بينما الاختبارات النفسية تحوي عينات صغيرة .
- تختلف نسبة المجموعات الدنيا والعليا بين الاختبارات النفسية والتربوية بسبب السمة المدروسة وحجم المجتمع فقد نعلم نسبة 27% حسب كيلي Kelly . أو بين 25 – 33% حسب كيورتن Cureton . أو بالمناصفة 50% .

خاتمة:

تم من خلال بحثنا التعريف بعملية تقنين الاختبارات النفسية والتربوية في البحث العلمي. وتوصلنا إلى تأكيد ضرورة القيام بهذه العملية في البحث العلمي؛ لما لها من أهمية في التحقق من صلاحية الاختبار كأداة للقياس والتشخيص فيما بعد سواء على المستوى النفسي أو التربوي. كما تمت الإشارة إلى بعض الاختلافات في عملية التقنين بين الاختبارات النفسية والتربوية. على اعتبار أن كل نوع يقيس سمة من سمات المفحوص.

CONCLUSION

Through our research we defined the process of Standardization psychological and educational tests in scientific research. And We have come to confirm the necessity of carrying out this process in scientific research; Because of its importance in verifying the validity of the test as a tool for measurement and diagnosis later, whether at the psychological or educational level.

We were also indicated Some differences in the Standardization process between psychological and educational tests. Assuming that Each type measures one of the characteristics of the Examined.

قائمة المراجع:

- 1- الداوودي، تامر. جانفي 2013. المكتبة الرياضية الشاملة. تاريخ الاسترداد 20 نوفمبر 2020، من المعايير والمستويات: <https://www.sport.ta4a.us/human-sciences/tests-measurements/579-standards-and-levels.html>
- 2- السيد، محمد أبو هاشم حسن. 2006. محاضرة الخصائص السيكمترية لأدوات القياس في البحوث النفسية والتربوية باستخدام SPSS. السعودية: كلية التربية. جامعة الملك سعود.
- 3- العساف، صالح حمد. 2000. المدخل إلى البحث في العلوم السلوكية. الرياض: مكتبة العبيكان.
- 4- أبو ناهية، صلاح الدين محمد. 1994. القياس التربوي. الاسكندرية: المكتبة الأنجلو المصرية.
- 5- اسماعيل، بشرى. 2004. المرجع في القياس النفسية. الاسكندرية: مكتبة الأنجلو المصرية.
- 6- بوقرة، عواطف. 2021. الوجيز في القياس النفسي. الجزائر: دار قريطا للنشر والتوزيع.
- 7- بوسنة، محمود. 2007. علم النفس القياسي - المبادئ الأساسية-. الجزائر: ديوان المطبوعات الجامعية.
- 8- زيدان، محمد. 1979. معاجم المصطلحات النفسية والتربوية. ط1. جدة: دار الشروق.
- 9- سعد، عبد الرحمن. 1998. القياس النفسي بين النظرية والتطبيق. ط3. القاهرة: دار الفكر العربي.
- 10- صفوت، فرج طه. 2007. القياس النفسي. ط1. القاهرة: مكتبة الأنجلو المصرية.
- 11- علام، صلاح الدين محمود. 2000. القياس والتقويم التربوي والنفسي. القاهرة: دار الفكر العربي.
- 12- معمري، البشير. 2012. أساسيات القياس النفسي وتصميم أدواته للطلاب والباحثين في علم النفس والتربية. الجزائر: دار الخلدونية للنشر والتوزيع.

13-مراد، صلاح أحمد . سلمان، أمين علي . 2005. الاختبارات والمقاييس في العلوم النفسية والتربوية. القاهرة: دار الفكر العربي.

Bibliography List :

- 1- Al-Assaf, Saleh Hamad. 2000. Introduction to Research in the Behavioral Sciences. Riyadh: Obeikan Library.
- 2- Abu Nahia, Salah al-Din Muhammad. 1994. Educational Measurement. Alexandria: The Anglo-Egyptian Library.
- 3- Allam, Salah El-Din Mahmoud.2000. Educational and psychological measurement and evaluation. Cairo: Arab Thought House.
- 4- AlDaoudi, Tamer. January 2013. The Comprehensive Mathematical Library. Retrieved November 20, 2020, from Standards and Levels: <https://www.sport.ta4a.us/human-sciences/tests-measurements/579-standards-and-levels.html>
- 5- Bouguerra,Awatif. 2021. Al-Wadjeez in Psychometrics. Algeria: Dar Quirta Publishing and Distribution.
- 6- Boussina, Mahmoud. 2007. Standard Psychology - Basic Principles -. Algeria: Diwan of University Press.
- 7- Groth, Marnat, G. 2009. Handbook of Psychological Assessment.5 th Ed.Canada: John Wiley and Sons.
- 8- Ismail, Bushra. 2004. The reference in psychometrics. Alexandria: The Anglo-Egyptian Library.
- 9- Maamriya, elBashir. 2012. The basics of psychometrics and the design of its tools for students and researchers in psychology and education. Algeria: Dar Al Khaldouniyah for Publishing and Distribution.
- 10- Murad, Salah Ahmed. Salman, Amin Ali. 2005. Tests and Measurements in Psychological and Educational Sciences. Cairo: Arab Thought House.
- 11- Saad, Abdel Rahman. 1998. Psychometrics between theory and practice. i 3. Cairo: Arab Thought House.
- 12- Said. Muhammad Abu Hashem Hassan. 2006. A lecture on the psychometric properties of measurement tools in psychological and educational research using SPSS. Saudi Arabia: College of Education. King Saud University.
- 13- Safwat, Faraj Taha. 2007. Psychometrics. i 1. Cairo: Anglo-Egyptian Library.
- 14- Zeidan, Muhammad. 1979. Dictionaries of psychological and educational terms. i 1. Jeddah: Dar Al-Shorouk.

Standardization of Tests in Psychological and Educational research

Bouguerra Awatif ¹

¹ M'sila University, Algeria

Email adress, awatif.bouguerra@univ-msila.dz

Abstract:

The current study sought to define the process of Standardization tests in psychological and educational research; As the most important step of psychological and educational measurement in scientific research. Because it confirms the validity of the measuring instrument used and verifies its validity and Reliability. It also sought to clarify the steps of the Standardization process. It concluded that there are some differences in this process between psychological and educational research, as it is not relied upon to measure the difficulty and ease of paragraphs in some psychological tests, except for those related to the study of mental abilities. Also, the effectiveness of subliminals and their indicators in psychological tests is not studied, unlike educational tests that depend on measuring academic achievement in most research studies. It also reached to confirm the importance of the step of verifying the level of measurement error in both psychological and educational studies because of its importance.

Keywords: Standardization of tests, Psychological research, Educational research