Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/110121

Internal Consistency Between Validity and Reliability in Academic Research: An Evaluation Study

Amel Mohamed Beichi[1,3], *Sakina Mhamed Tchalabi*[4] *and Najeh Rajeh Alsalhi* [1,2,3,*]

[1]College of Humanities and Sciences, Ajman University, Ajman, UAE
[2]Nonlinear Dynamics Research Center (NDRC), Ajman University, Ajman, UAE
[3]Humanities and Social Sciences Research Center (HSSRC), Ajman University, Ajman, UAE
[4]College of Humanities and Social Sciences, Blida University, Algeria

Received: 22 Jul. 2021, Revised: 2 Sep. 2021, Accepted: 16 Oct. 2021.
Published online: 1 Jan. 2022.

Abstract:

This study is intended to discuss the dialectic of internal consistency between reliability and validity through the presenting of opinions of experts and psychometric experts and to then evaluate a sample of Arab and foreign studies that sought to verify the validity of psychological tests before using. The sample of study consisted of (83) tests, distributed among doctoral and master's theses and articles, in more than (20) countries. The study used the descriptive evaluation method because it is related to the nature and objectives of the study. The analysis was performed using an evaluation network tool prepared by researchers. The findings revealed that the internal consistency is the most used in checking and confirming the validity, whereas, Cronbach's alpha is the most widely used measure of reliability.

Keywords: Internal consistency, Validity, Reliability, Academic Research.

*Corresponding author e-mail: n.alsalhi@ajman.ac.ae

© 2022 NSP
Natural Sciences Publishing Cor.

الاتساق الداخلي بين الصدق والثبات في البحوث الأكاديمية: دراسة تقييمية

آمال محمد بايشي[1،3] سكينة محمد تشلابي [4] ناجح راجح الصالحي [1،2،3]

1 كلية الإنسانيات والعلوم، جامعة عجمان، عجمان، الإمارات العربية المتحدة.

2 مركز أبحاث الديناميات غير الخطية، جامعة عجمان، عجمان، الإمارات العربية المتحدة.

3 مركز أبحاث العلوم الإنسانية والاجتماعية (HSSRC) ، جامعة عجمان ، عجمان ، الإمارات العربية المتحدة.

4 كلية العلوم الإنسانية والاجتماعية، جامعة البليدة، الجزائر.

الملخص

تهدف هذه الدراسة إلى مناقشة جدلية الاتساق الداخلي بين الصدق والثبات، وذلك من خلال عرض آراء الخبراء والمختصين في مجال القياس النفسي، ومن ثم تقييم عينة من الدراسات العربية والأجنبية التي عمدت إلى التحقق من صلاحية الاختبارات النفسية قبل استخدامها، وتكونت العينة من ثلاثة وثمانين (83) اختبارًا، موزعة على رسائل الدكتوراه والماجستير والمقالات، في أكثر من عشرين (20) بلدًا، وقد تمّ استخدام المنهج الوصفي التقويمي لملاءمته طبيعة الدراسة. وكذا الاعتماد على شبكة تقييم أعدّها فريق من الباحثين، وقد أظهرت النتائج أن ما يطلق عليه صدق الاتساق الداخلي هو الأكثر استخداما في التحقق من الصدق، في حين أن معامل "ألفا كرونباخ" هو الإجراء الأكثر استخداما في التحقق من الثبات.

الكلمات المفتاحية: الاتساق الداخلي، الصدق، الثبات، البحث الأكاديمي.

1 المقدمة

يضطر الباحث أو الفاحص بصفة عامة إلى التحقق من صدق وثبات الاختبار النفسي سواء عند بنائه أو عند استخدامه جاهزا، خاصة إذا كان هذا الأخير غير مكيف للبيئة الجديدة أو تم تغيير مجتمع البحث عن المجتمع الأصلي للاختبار وغيرها من الضوابط، فكما أشار[1] أن عملية التحقق من صلاحية الاختبار هي بمثابة مسؤولية مشتركة بين مطوّر الاختبار ومستخدمه. فالصدق ليس أمرا مطلقا حسب كل منThorndike وHagen بل يختلف من اختبار لآخر، بحيث لا نستطيع أن نقول أن الاختبار إما صادق أو غير صادق بل نقول إنه صادق بقدر أو بدرجة ما، ووفقا لما أشار إليه [2]، والصدق تقدير لما إذا كان الاختبار يقيس ما نريد أن نقيسه به، وكل ما نريد أن نقيسه به، ولا شيء غير ما نريد أن نقيسه به، وينطبق الأمر نفسه في الثبات، فالصدق والثبات صفتان لا ترتبطان بالاختبار في حد ذاته بل ترتبط بالنتائج. لكن رغم هذا فإن العديد من الباحثين يكتفون بالخصائص السيكومترية الواردة للاختبارات في البيئات الأصلية لها، رغم تغير العينة والغرض من الاختبار وغيرها، في حين يعمل البعض الآخر على محاولة استخراج دلالات صدق وثبات جديدة للاختبارات كإجراء إحصائي روتيني لإعادة التحقق من صلاحية الاختبارات النفسية، ويوضح [3] بأنه لا يكاد يخلو بحث أو رسالة من استعمال صدق الاتساق الداخلي حيث يتم ذلك من خلال إيجاد معاملات الارتباط بين كل فقرة بُعدها، وأحيانا بالمقياس ككل، وارتباط الأبعاد بالاختبار ككل، ويتم الحكم على اتساق الفقرة بناءً على الدلالة الإحصائية بحيث تستبعد كل فقرة غير دالة إحصائيا، وهنا قد يجد الباحث نفسه في تناقض بين نتائج صدق الاتساق الداخلي والتحليل المنطقي لبنود الاختبار، فعند دراسة الارتباط بين الفقرات وتبيّن أن أحد البنود أو عدة بنود غير دالة فإنه وجب حذفها انطلاقا من المنطق الإحصائي، في المقابل فإن المتفحص لتلك البنود قد يجدها تجسد وتعكس السمة المدروسة أكثر من البنود الأخرى التي قبلت إحصائيا، وهنا يقع الباحث في إشكال محرج هل يحذف الفقرة وهي مهمة بالنسبة للسمة أم يبقيها وهي غير دالة. وما عادا هذا الإشكال يطرح الاتساق الداخلي العديد من الإشكالات التي سنحاول الإجابة عنها، مثل هل يعتبر الاتساق الداخلي صدقا أم ثبات؟ هل يعتبر حقا من أكثر الطرق استخداما عند الباحثين؟ ومعاملات الارتباط هل هي مقبولة؟ هل يستخدم الباحثون الاتساق الداخلي كإجراء للتحقق من الصدق والثبات في آن واحد؟ وأخيرا هل يعتبر الاتساق الداخلي الإجراء الأكثر شيوعا واستعمالا للتحقق من الثبات؟

2 أهداف الدراسة

تسعى هذه الدراسة من خلال تقييم عينة من الدراسات العربية والأجنبية إلى ما يلي:

- مناقشة جدلية الاتساق الداخلي بين الصدق والثبات.
- تحديد الطريقة الأكثر استعمالا في التحقق من صدق الاختبارات النفسية.
- الكشف عن الأخطاء التي يقع فيها الباحثون عند استخدام الاتساق الداخلي كإجراء للتحقق من الصدق أو الثبات.
- التعرف على نوع الثبات الأكثر شيوعا عند التحقق من صلاحية الاختبارات النفسية.

© 2022 NSP
Natural Sciences Publishing Cor.

3 منهج الدراسة

أعتمد الباحثون في دراستهم الحالية على المنهج الوصفي التقويمي، لملائمته لمتطلبات الدراسة وأهدافها، ويعتمد هذا المنهج في وصفه على التكرارات والنسب المئوية من أجل وصف وتقييم عينة من الدراسات العربية والأجنبية التي تم فيها التحقق من صدق وثبات الاختبارات النفسية قبل استخدامها.

3.1 مصطلحات الدراسة

قبل التطرق إلى الجدل القائم حول الاتساق الداخلي فيما إذا كان صدقا أم ثباتًا ومحاولة مناقشته وجب علينا الانطلاق من تعريف الصدق والثبات ومن ثم الاتساق الداخلي، ليكون في الأخير انتساب الاتساق الداخلي إلى أحدهما دون الآخر منطقيا ومفهوما ومدروسا.

3.2 الصدق

عرّفه [4] على أنه حكم تقييمي شامل للدرجة التي يُتحصل عليها من الاختبار والتي تؤيّدها الأدلة التجريبية والنظرية، أي من حيث ملائمة الأدلة وكفايتها، اعتمادا إلى درجات الاختبارات وغيرها من التقييمات. كذلك عرّفه [5] على أنه العملية التي عن طريقها يجمع مستخدم الاختبار أو مطوره، الدلائل من درجات الاختبار والتي يدعم من خلالها استدلالاته. وعند التحقق من الصدق أو الصلاحية، يجب تحديد الاستدلال المطلوب، لكي يتم جمع الأدلة المناسبة لهذا النوع من الاستدلال. وعليه فالصدق يعتمد على الأدلة النظرية والتجريبية أكثر من اعتماده على المنطق الإحصائي على عكس الثبات الذي يستند بشكل كبير إلى الأدلة الإحصائية، كذلك أضاف كرونباخ فكرة مهمة فيما يتعلق بالصدق، فليست كل الأدلة صالحة لجميع الاختبارات، وهذا بالطبع راجع لاختلاف الغرض من الاختبار فلكل اختبار غرضه الخاص وبالتالي لكل اختبار أدلة معينة تثبت صلاحيته لذلك الغرض.

3.3 الثبات

يُشير الثبات حسب [6] إلى اتساق الدرجات التي يحصل عليها نفس الأشخاص عندما يعاد اختبارهم بنفس الاختبار في وقتين مختلفين، أو بمجموعتين مختلفتين من المفردات المتكافئة، أو في ضوء ظروف اختباريه لمتغير آخر. وينطوي هذا المفهوم للثبات على حساب خطأ القياس لدرجة واحدة، والتي يمكن بواسطتها التنبؤ بمدى التذبذب المحتمل حدوثه في درجة واحدة للفرد نتيجة لعوامل صدفة غير موائمة أو غير معلومة، كما تم تعريف الثبات حسب وثيقة معايير القياس النفسي [7] على أنه لا يعتبر خاصية من خصائص الاختبار في حد ذاته وإنما هو مرتبط بالدرجات والنتائج. وهو الأمر نفسه الذي ينطبق على الصدق في هذه النقطة فالصدق والثبات غير مرتبطين بالاختبار في حد ذاته فلا نقول أن الاختبار صادق أو ثابت بل نقول أن الاختبار صادق أو ثابت في ما يقيسه، أي أن الصدق والثبات متعلقين بالنتائج.

3.4 الاتساق الداخلي

إن مصطلح الاتساق الداخلي يرمي إلى خاصيتين متمايزتين، فتارة يشار به إلى صدق الاتساق الداخلي (Internal consistency validity)، وتارة يشار به إلى الثبات باستخدام معامل ألفا كرونباخ، إذن فالاتساق داخلي قد يعتبر صدقا وقد يعتبر ثبات لذلك عرفه [3] على أنه إجراء يخدم صدق التكوين الفرضي، ولا يعدو أن يكون مرادفا آخر لقياس فكرة الاتساق التي تشكل لحمة مفهوم الثبات، فنجد أن الباحث من غير وعي منه يقيس الثبات مرتين: مرة تحت مسمى الثبات بـ" معادلة ألفا كرونباخ"، ومرة أخرى تحت مسمى صدق الاتساق الداخلي عن طريق ربط الفقرات ببُعدها أو بالاختبار ككل. ويشير [6] إلى أنه من أجل فهم أسلوب تقدير الاتساق الداخلي كصدق للاختبار يتعين علينا التمييز بين طريقتين لتطبيق هذا الأسلوب والمجالات التي تستخدم فيها كل طريقة:

3.4.1 الطريقة الأولى: هي الارتباط بين البند والدرجة الكلية، وتفسر معاملات الارتباط الناتجة عن هذا الأسلوب بالطريقة نفسها عندما نستخدم الارتباط بين البند والدرجة الكلية لحساب مؤشر الصعوبة في مجال مقاييس القدرات فالارتباط الموجب يشير إلى نجاح البند في التمييز بين أصحاب الأداء الجيد على الاختبار وأصحاب الأداء الضعيف، كما يشير أيضا وهذا هو الأمر الأكثر أهمية إلى أن البند يقيس الشيء نفسه الذي يقيسه الاختبار ككل ولكنه لا يتضمن معلومة محدَّدة حول ما هو هذا الشيء نفسه الذي يقيسه الاختبار ككل.

3.4.2 الطريقة الثانية: هي الارتباطات الداخلية بين البنود، ويساعد هذا على فهم القدرة التمييزية للبنود، فمن المتوقع أن يظهر البند صاحب الارتباط المرتفع بالدرجة الكلية ارتباطات مرتفعة بأغلب البنود على الاختبار نفسه، وقد يظهر فحص الارتباطات بين البنود السبب في كون بعض البنود دون غيرها تفشل في التمييز بين طرفي السمة في الأداء. وفي كلتا الحالتين لا يمكن الخروج بدلالات حول صدق ما يقيسه الاختبار من خلال المعلومات المتعلقة بالارتباط بين البنود والدرجة الكلية أو من الارتباطات البينية بين البنود.

4. جدلية الاتساق الداخلي صدق أم ثبات؟

إن موضوع الاتساق الداخلي يفتح باباً واسعاً للنقاش حوله، فإشكالية اشتراك كل من الصدق والثبات في هذا الإجراء جعلت منه محط اصطدام دائم، مابين اعتباره ثباتا وفقط، أو اعتباره إجراء صالح حتى لإثبات صدق الاختبار، وفقا لما ورد عن [3] قد نعتمد على الأدلة أو البيّنات القائمة على الاتساق الداخلي كدليل على صدق الاختبار، في الوقت الذي قد نستعمل نفس الدليل للحكم على ثبات المقياس القائم على الاتساق الداخلي باستخدام معامل ألفا كرونباخ، أي استعمال الاتساق الداخلي لقياس الثبات واستعمال نفس الاتساق الداخلي لقياس الصدق، وبهذا نكون قد استعملنا نفس الحجة مثارة تحت مظلة الثبات، وتارة أخرى تحت مظلة الصدق. أضف إلى ذلك أن الشواهد أو الأدلة أو البيّنات المستعملة للدلالة على الصدق هي أدلة قائمة على الاتساق الداخلي للمقياس، علماً بأننا نستعمل نفس الدليل للحكم على ثبات المقياس القائم على الاتساق الداخلي باستخدام معامل(ألفا كرونباخ) للاتساق الداخلي[استعمال الاتساق الداخلي لقياس الثبات واستعمال نفس الاتساق الداخلي لقياس الصدق]، وبهذا نكون قد استعملنا نفس الحجة تارة تحت مظلة الثبات، وتارة أخرى تحت مظلة الصدق، وفي ذلك تكرار لنفس الدليل لغرضين مختلفين، علماً بأن الحجة القائمة على الاتساق الداخلي حجة تنتمي أصلا إلى الثبات.

© 2022 NSP
Natural Sciences Publishing Cor.

ومن الواضح أنه وفقا لما أشار إليه [2] أن الاتساق الداخلي سواءً كان للمقاييس الفرعية أو للبنود فإنه أساسا يقيس التجانس ولأنه يحدّد خصائص المجال السلوكي أو للسمة التي يقيسها الاختبار، فإن درجة تجانس الاختبار تتعلق بقدر ما بصدق التكوين، وتضيف كل من [6] أنه على الرغم من ذلك فإن الإضافة التي توفرها بيانات الاتساق الداخلي لصدق الاختبار محدودة ففي ضوء غياب بيانات من خارج الاختبار نفسه فلن نعرف إلا القليل عن ما يقيسه الاختبار. وعلى هذا فمن الأفضل الاكتفاء باستخدام أساليب الاتساق الداخلي في مجالي تحليل البنود لأغراض التمييز والصعوبة أو لحساب ثبات الاختبار.

كما أوضح [6] أن صدق التكوين الفرضي يقسم إلى الصدق الذاتي للاختبار وصدق التفسير، فقِيم الارتباط الداخلي بين مفردات الاختبار، والقيم المرتفعة للارتباط بين درجات الاختبار ومتغيرات خارجية تقدم أدلة على أن الاختبار يقيس شيئا ما بطريقة منتظمة، أما محتوى المفردات وطبيعة العلاقات القائمة بينها، وطبيعة المتغيرات الخارجية التي تتباين في درجة ارتباطها بالاختبار فتقدِّم أدلة على التفسير المقترح. أي أن الأدلة الأولى تُفيد في إضفاء معنى سيكومترية على الاختبار، أما الأدلة الثانية فتضفي على الاختبار معنى سيكولوجي أو تربوي، وكلاهما أمراً ضرورياً[6]. وبالتالي فكل من أنسنازي ولوفينجر اتفقتا أن الاتساق الداخلي حتى وإن ضُمّ تحت جناح الصدق وبالضبط صدق التكوين الفرضي، إلاَّ أنه غير كافٍ أبداً بحيث أنه يضيف معنى سيكومترية وفقط، فالاتساق الداخلي يعني أن البنود جميعها تقيس شيئا مشتركا أو تقيس ما يقيسه البعد أو المحور بصفة عامة، لكن هذا لا يعني أن البنود تقيس ما وضعت لقياسه، أو تقيس ما وضعت لقياسه وفقط، ولا يدل كذلك على أن البنود تقيس أمراً إضافيا، وبالتالي فكيف يمكن الاكتفاء به كدليل على صدق الاختبار. وغير بعيد عن هذه الفكرة عبرت [8] أنه ليس من المناسب اعتبار الاتساق الداخلي مقياساً للصدق لأنه لا يخبرك ما إذا كانت الفقرات تقيس البناء الذي يراد قياسه. بل إنه يخبرك فقط ما إذا كان الاختبار يقيس شيئا مشتركا، ولكن ما هو ذلك الشيء؛ لا يزال غير معروف. كذلك كما ذكرنا سابقا في المقدمة يطرح الاتساق الداخلي عند استخدامه في الصدق بخاصة إشكالا يصادف العديد من الباحثين، ألا وهو الدلالة الإحصائية في مقابل الدلالة النظرية، فيجد الباحث نفسه في صراع بين قبول المنطق الإحصائي أو المنطق النظري. وقد أرجع [3] السبب في هذا التناقض إلى أن الدلالة النظرية لفقرة ما لا توازي بالضرورة الدلالة الإحصائية لها، لكون الدلالة الإحصائية تتأثر بعوامل كثيرة لا علاقة لها بالدلالة النظرية للفقرة، منها مدى تجانس الإجابات، أو تباينها، ومنها أيضا حجم العينة، وقوة الاختبار الإحصائي ولكن من أهمها إشكالية تزييف الاستجابة من طرف المفحوصين ومدى جديتهم في الإجابة على بنود المقياس أو الاختبار، بمعنى أن التناقض بين نتائج التحليل الإحصائي والتحليل المنطقي للباحث وباحثين آخرين يمكن إرجاعه إلى المفحوصين، كما اقترح [9] الاعتماد في تقنيين الاختبارات على العينات المقصودة والتي تقبل فعلا التعامل مع الباحث وتساعد في الإجابة بموضوعية. أيضا تفترض الاستنتاجات المستندة إلى تحليل الاتساق الداخلي أن الأداة المعنية، التي نحاول إثبات صلاحيتها، تقيس بُنية أحادية البعد، ومع ذلك يمكن للأداة في الحقيقة قياس عدد من التركيبات أو البنيات الفرعية أي الأبعاد [10]، ولذلك وجب هنا التعامل مع كل بعد على أنه محور مستقل في حدِّ ذاته. إذن، واختصارا للنقاش السابق، فإن الاتساق الداخلي للبنود لا تعتبره كل من [6] و [3] وغيرهم صدقاً ولكن يمكن استغلاله في الصدق لإثبات التكوين الفرضي للبنود وبأنها تقيس شيئا مشتركا ولكن لا يمكن الاكتفاء به أبداً للدلالة على صدق الاختبار بل وجب دائما أن تصاحبه أدلة أخرى تثبت صدق الاختبار، وكمثال للتبسيط أكثر إذا أراد باحث ما التحقق من صدق اختبار يقيس الاحتراق النفسي مثلا واستعان لإثبات صدق اختباره بالعلاقة بين البنود والدرجة الكلية أو بين البنود فيما بينها وتم إثبات العلاقة ودلالتها فهل هذا يعني أن الاختبار يقيس الاحتراق النفسي؟ بالطبع لا، فقط يدل على أن البنود تقيس متغير مشترك وقد يكون هذا المتغير ضغوط نفسية أو اضطراب بعد الصدمة أو متغير آخر. إذن فالاتساق الداخلي يدل فقط على أن الاختبار يقيس شيئا مشتركا ولا يدل على أنه يقيس ما وضع لقياسه. وفي الأخير أضاف [6] أمرا آخر بأن لمستخدم الاختبار القرار النهائي بشأن صلاحية وفائدة اختبار أو مقياس ما، وجب عليه وصف أساس هذا القرار وتبريره في قسم الإجراءات من خطة البحث الخاصة به. أي أنه أكّد على أن القرار راجع إلى مستخدم الاختبار في اختياره للأدلة المناسبة للتحقق من صدق الاختبار بما يتوافق والغرض منه آخذا بذلك جميع الاعتبارات السابقة، ولكن يبقى كل ذلك مرهونا بتبريره لذلك الاختيار تبريراً منطقيا.

5. منهجية البحث

5.1 إستراتيجية البحث

أُجري البحث على عينة الدراسة والتي تتمثل في الدراسات والمقالات سواءً العربية أو الأجنبية من خلال قواعد البيانات الإلكترونية، وبالنسبة للدراسات العربية تم اعتماد قاعدة المعرفة وASKZAD وقاعدة بيانات المنظومة. أما في ما يخص الدراسات الأجنبية فقد تمّ الاعتماد على قاعدة Science Direct و Springer Link وذلك باستخدام مزيج من عناوين الموضوعات التي تستخدم فيها عادة الاختبارات النفسية ومن هذه العناوين مثلا: (اختبارات الشخصية، الاكتئاب، الضغط النفسي.....الخ)

5.2 عينة الدراسة

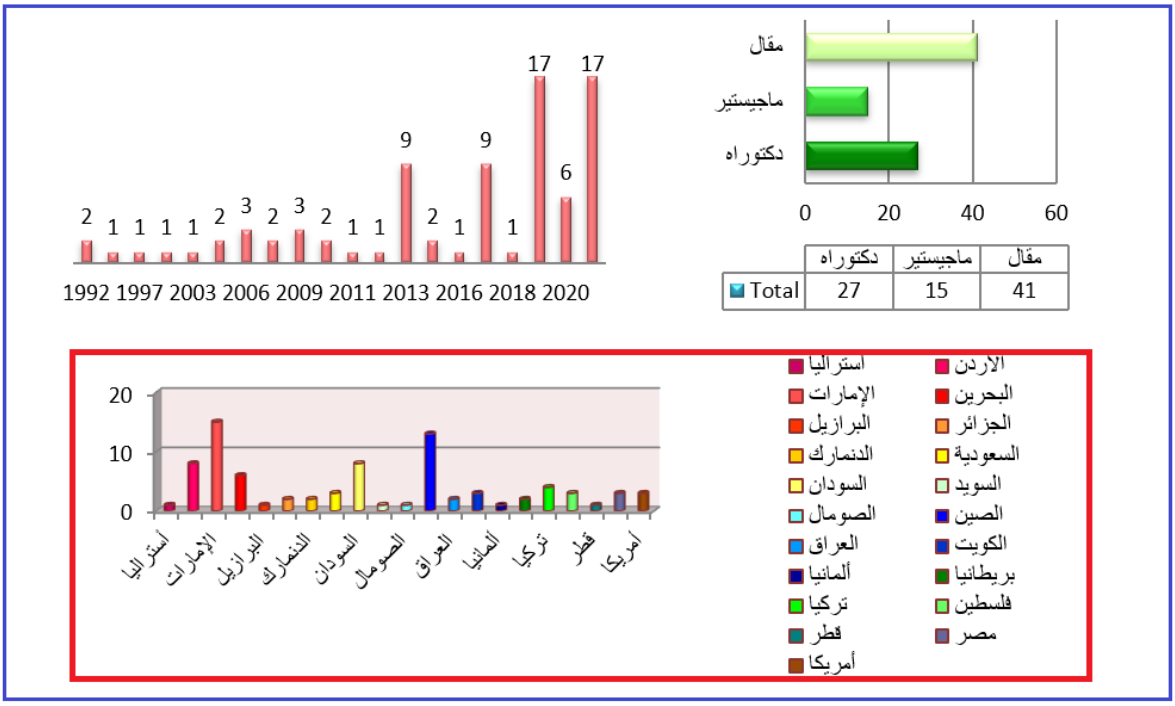
تمّ فرز عناوين الدراسات والمقالات وملخصاتها والتحقق منها. واعتماد الدراسات التي تستوفي الشروط فقط، أما المواد التي لا تستوفي معايير الاشتمال فقد استبعدت استناداً إلى عناوينها أو مضمونها، وكانت معايير القبول في الدراسة هي:

أولا، الاعتماد على اختبار نفسي.

ثانيا، إعادة التحقق من صلاحية الاختبار سواء تمّ بناءه أو استخدام اختبار جاهز.

ثالثا، استبعاد الدراسات التي اعتمدت على التحقق من صلاحية الاختبارات نفسية دون ذكر الكيفية والتفاصيل. وقد تمّ في الأخير تقييم (83) اختباراً نفسياً، منها (28) اختباراً أجنبياً و (55) اختباراً عربياً من مقالات ورسائل دكتوراه وماجستير موزعة على العديد من البلدان وكذا السنوات كما سنوضحه في الأشكال التالية:

© 2022 NSP
Natural Sciences Publishing Cor.



الشكل (1): عينة الدراسة.

.

كما هو واضح من الأشكال السابقة فقد تمّ الاعتماد على عينة غير متجانسة، فقد تعمد فريق البحث إلى محاولة سحب العينة من عدة بلدان مختلفة وفي سنوات متفاوتة، كما هو واضح في الشكل أعلاه، كمحاولة للتوسع في تباين العينة، بحيث تراوحت سنوات المناقشة للرسائل والنشر بالنسبة للمقالات بين 1992 و2021، أما البلدان فتوزعت على أكثر من (20) بلد عربي وأجنبي. وتمايزت الدراسات بين رسائل دكتوراه ورسائل الماجستير والمقالات.

5.3 الأساليب الإحصائية

بهدف الإجابة عن تساؤلات الدراسة تمّ الاستعانة ببرنامج الحزمة الإحصائية (SPSS) لاستخدام الإحصاء الوصفي من تكرارات ونسب مئوية.

5.4. أداة الدراسة

تمّ استخراج معلومات من الدراسات التي تستوفي معايير القبول باستخدام شبكة تقييم منظمة ومسبقة التصميم. وشملت هذه الشبكة: بيانات عامة مثل اسم صاحب المقال أو الدراسة وسنة نشرها وعنوان الدراسة واسم الاختبار المستخدم، وكذلك البلد والمشرف بالنسبة لرسائل الماجيستر والدكتوراه. كذلك تضمّنت الشبكة أسئلة حول أصل الاختبار هل هو من نفس البيئة أو بيئة مختلفة، وبيانات خاصة حول الصدق والثبات.

6. عرض نتائج الدراسة

6.1 نتائج الإجابة عن التساؤل الأول: ما هي أكثر أنواع الصدق استخداماً في الدراسات الأكاديمية العربية والأجنبية؟

للإجابة عن هذا التساؤل قمنا بتحليل 83 اختبار تمّ التحقق من صلاحيته، وكانت النتائج كما هي مبينة في الجدول الآتي:

الجدول (2): أنواع الصدق المستخدمة

أنواع الصدق	التكرار
عرض صدق الاختبار الأصلي	5
الصدق التمييزي	1
الصدق الذاتي	1
الصدق الذاتي، صدق الاتساق الداخلي	1

© 2022 NSP
Natural Sciences Publishing Cor.

الصدق المحكي	1
الصدق المحكي، صدق الاتساق الداخلي	4
تم التحقق من الثبات فقط	20
صدق الاتساق الداخلي	25
صدق الاتساق الداخلي، الصدق التمييزي	1
صدق الاتساق الداخلي, الصدق التمييزي، التحليل العاملي	1
صدق البناء	1
صدق المحتوى، الصدق الظاهري، الصدق التمييزي	1
صدق المحكمين	13
صدق المحكمين، الصدق المحكي، الصدق التمييزي	1
صدق المحكمين، صدق الاتساق الداخلي	2
صدق المحكمين، صدق المقارنة الطرفية، صدق الاتساق الداخلي	1
صدق المحكمين، صدق المقارنة الطرفية، صدق البناء، صدق الاتساق الداخلي	2
صدق المقارنة الطرفية، صدق البناء، الصدق المحكي	1
عرض الصدق في دراسة أخرى	1
العدد الإجمالي	**83**

المصدر: من إعداد فريق البحث

كما هو واضح من الجدول تمّ عرض أنواع الصدق كما هي مستخدمة في الدراسات، أي لم نفصل بين الأنواع، بل تمّ ذكر الأنواع التي استخدمت في الدراسات سواء نوع واحد أو عدة أنواع، فمن المطلوب أن لا يكتفي الباحث بنوع واحد من الصدق في إثبات صلاحية اختباره بل وجب عليه جمع العديد من الأدلة التي تبرهن صلاحية الاختبار للغرض المرجو منه، وبالتالي قمنا بعرض كل الأنواع التي استخدمها الباحث في دراسته، وكما هو ملاحظ فإن أكثر الأنواع استخداما هو صدق الاتساق الداخلي في 25 دراسة وصدق المحكمين في 13دراسة وكلاهما استخدم بمفرده، على سبيل المثال لا الحصر دراسة [11] التي اكتفت بصدق المحكمين، أما [12] ودراسة [13] مثلا فاكتفت بصدق الاتساق الداخلي، وهو من الأخطاء الشائعة في الدراسات الأكاديمية بحيث لا يمكن الاكتفاء بصدق المحكمين أو صدق الاتساق الداخلي، فالأول يعتبر فقط إجراء من إجراءات صدق المحتوى والثاني اختلف فيه إذ يعتبر أصلا صدقا، واتفق أنه في حالة اعتبرناه كذلك وجب تدعيمه بأدلة أخرى كما هو الحال في العديد من الدراسات الموضحة في الجدول التي استخدمت عدة إجراءات وأدلة في التحقق من صدق الاختبار الواحد وهو ما نثمنه ونقدره.

6.2 *نتائج الإجابة عن التساؤل الثاني:* هل يعتبر الاتساق الداخلي من أكثر الإجراءات اعتمادا في التحقق من الصدق عند الباحثين؟

الجدول (3): تكرارات صدق الاتساق الداخلي

صدق الاتساق الداخلي	التكرار
الصدق الذاتي، صدق الاتساق الداخلي	**1**
الصدق المحكي، صدق الاتساق الداخلي	**4**
صدق الاتساق الداخلي	**25**
صدق الاتساق الداخلي، الصدق التمييزي	**1**
صدق الاتساق الداخلي، الصدق التمييزي، التحليل العاملي	**1**
صدق المحكمين، صدق الاتساق الداخلي	**2**
صدق المحكمين، صدق المقارنة الطرفية، صدق الاتساق الداخلي	**3**
العدد الإجمالي	**37**

المصدر: من إعداد فريق البحث

كما هو متوقع فالاتساق الداخلي يعتبر من الإجراءات الأكثر استخداما في الدراسات العربية و الأجنبية سواء رافقه الباحث بإجراء آخر، أو اكتفى به كدليل على صدق اختباره، فقد تمّ اعتماده في التحقق من صدق 37 اختبار من مجموع 83 اختبار، في 25 اختبار تم الاعتماد على الاتساق الداخلي فقط، في حين تم اعتماده مصاحبا لإجراءات أخرى في 12 اختبار، مثل دراسة [14] ودراسة [15]، وهو أمر يستدعي الوقوف عنده، فالاتساق الداخلي يكشف لنا إذ ما كانت البنود والمحور أو البنود والبعد تقيس أمرا مشتركا فقط، ولا يكشف لنا ماهية هذا المتغير الذي يقيسه الاختبار، وبالتالي فإننا نناقض أهم أساسيات الصدق ألا وهو أن

© 2022 NSP
Natural Sciences Publishing Cor.

يقيس الاختبار ما وضع لقياس وما وضع لقياسه وفقط، وهنا الاتساق الداخلي يثبت لنا جزئية منه فقط. فكيف لـ 25 دراسة من أصل 37 أن تكتفي بالصدق فقط كدليل على صلاحية الاختبار؟ في حين أنه لا يكشف لنا أصلا عما يقيسه الاختبار.

6.3 نتائج الإجابة عن التساؤل الثالث: هل معامل صدق الاتساق الداخلي مقبول؟

الجدول (4): تكرارات صدق الاتساق الداخلي

معامل الصدق مقبول؟	التكرار
نعم	15
لا	22
العدد الإجمالي	**37**

المصدر: من إعداد فريق البحث

بما أن الاتساق الداخلي يعتمد على معامل الارتباط بين البنود والدرجة الكلية سواء للمحور أو البعد، وبين البنود فيما بينها، فإنه من الأخطاء الأخرى التي يقع فيها الباحثون ما عادا استخدام الاتساق الداخلي كإجراء للصدق منعزلا، هي قبول جميع القيم معامل الارتباط التي يصل إليها الباحث والذي يضعنا أمام تساؤل مهم ما هو الهدف من التحقق من صدق الاختبار إذا كان بإمكاننا قبول جميع القيم، صغيرة جدا كانت أم كبيرة؟ خاصة وأن أغلبية الباحثين لا يبرّرون سبب قبول تلك القيم المنخفضة والأمر من ذلك الغير دالة أحيانا، فنادرا ما يلجأ الباحث إلى حذف البند المنخفض، بل يكتفي فقط بالإشارة إلى أنه بند ذو ارتباط ضعيف مثل دراسة [16]. وأيضا دراسة [11] ودراسة [17] التي لم تتجاوز معاملاتها 0.20 إلا أن الباحثين احتفظوا بتلك البنود الضعيفة والاختبارات، وفي أحسن الأحوال كان بعض الباحثين يقوم بحذف البنود الغير دالة فقط مثل دراسة [15] والبعض الآخر قرّر الاحتفاظ حتى ولو كانت غير دالة مثل دراسة [13] في حين لم يتطرق بعض الباحثين إلى ذكر مستوى الدلالة بل اكتفوا بذكر قيمة معامل الارتباط مثل دراسة [18] والذي قرّر أيضا حذف البنود الضعيفة والتي قدرت قيمتها بأقل من 0.25 والاحتفاظ بما هو أعلى من تلك القيمة.

6.4 نتائج الإجابة عن التساؤل الرابع: ما هي أنواع الثبات الأكثر شيوعا في التحقق من صلاحية الاختبار؟ وجاء هذا التساؤل للإشارة إلى الخطأ الذي يقع فيه بعض الباحثين عند استخدام الاتساق الداخلي كإجراء من إجراءات الصدق، وفي ذات الوقت استخدام ألفا كرونباخ كإجراء من إجراءات الثبات، وبالتالي يجد الباحث نفسه تحقق من صلاحية الاختبار النفسي بالطريقة ذاتها في الصدق والثبات إذ استخدم الاتساق الداخلي في الحالتين دون إدراك منه بهذا الخطأ.

الجدول (5): تكرارات صدق الاتساق الداخلي

أنواع الثبات	التكرار
اختبار ت	1
ألفا كرونباخ	64
بيرسون	13
عرض الثبات في الأصلية	5
العدد الإجمالي	**83**

المصدر: من إعداد فريق البحث

كما هو متوقع فإن ألفا كرونباخ هو الإجراء الأكثر شيوعا واستخداما في الدراسات العربية والأجنبية عند التحقق من الثبات، فقد اعتمد على معامل ألفا كرونباخ في التحقق من ثبات 64 اختبار من أصل 83 اختبار. مثل دراسة [19] ودراسة [20]. أما عند التحقق من عدد الاختبارات التي استخدم فيها الاتساق الداخلي كإجراء صدق أي بين البنود والدرجة الكلية أو بين البنود فيما بينها، وكإجراء للثبات باستخدام ألفا كرونباخ، وجد فريق البحث أن 22 اختبار تمّ التحقق من صدقه وثباته باستخدام نفس الإجراء فقط بطريقتين مختلفتين. مثل دراسة [21] ودراسة [22].

7. مناقشة النتائج

إن الحديث عن صلاحية الاختبارات النفسية يفتح المجال لدرجة يصعب معها تحديد الأهداف والاكتفاء بها، فكل هدف وكل تساؤل كان يأخذنا إلى هدف وتساؤل آخر، فعند محاولتنا الإجابة عن التساؤل الأول الذي كان يهتم بأكثر طرق القياس استخداما في الدراسات الأكاديمية، وجدنا أن الاتساق الداخلي هو أكثرها تكرارا يليه مباشرة صدق المحكمين، والذي بدوره يقودنا إلى العديد من التساؤلات الأخرى منها: هل يعتبر صدق المحكمين صدقاً حقا؟ والذي يعد خطأ شائعا هو الآخر، فصدق المحكمين هو إجراء فقط من إجراءات صدق المحتوى شأنه شأن الصدق الظاهري، وحتى وإن كان هذا الإجراء مهماً إلاّ أنه يبقى غير كاف، بالإضافة إلى ذلك يقع الباحثون في العديد من التجاوزات عند استخدامه مثل عدم احتساب نسبة الاتفاق بين المحكمين أو الاختيار غير الصائب للمحكمين وغيرها وهذا ما أشارت إليه أيضا واتفقت معه دراسة [23] ودراسة [24]. أما فيما يخص الاتساق الداخلي وهو الإجراء الأكثر استخداما حسب نتائج الدراسة الحالية والعديد من الدراسات الأخرى مثل دراسة [25] فـ27 من أصل 29 دراسة اعتمدت الاتساق الداخلي، كذلك دراسة [26] التي توصلت إلى أن جميع الدراسات التي قامت بتحليلها استخدمت الاتساق الداخلي، وأيضا اشتركت هذه الدراسة مع الدراسة الحالية في نقطة أخرى، وهي ضعف معاملات الارتباط أو حتى أحيانا عدم دلالتها بحيث 22 من أصل 37 اختبار كانت معاملات الارتباط فيه غير مقبولة، وقد أشارت دراسة [26] و دراسة [27] ودراسة [28] إلى وجود معاملات دون المستوى هي الأخرى. أيضا توصلت نتائج هذه الدراسة إلى أن معامل ألفا كرونباخ هو الأكثر استخداما في التحقق من الثبات، فـ 64 اختباراً من أصل 83 تمّ

© 2022 NSP
Natural Sciences Publishing Cor.

التحقق منه باستخدام ألفا كرونباخ، فهو الأكثر تداولا خاصة في الدراسات الأجنبية فقد اعتمدت كلها على ألفا كرونباخ وفقط. وقد توصلت دراسة كل من [26] و [29] وكذلك دراسة [30] إلى نفس النتائج واستهجنت قيام الباحثين بهذا. فعلى الرغم من أن هذا الإجراء هو الأكثر شيوعا، إلاّ أنه تم انتقاده لوجود العديد من العيوب فيه. على سبيل المثال يعتبر ألفا كرونباخ غير مناسب لمقياس متعدّد الأبعاد وبالتالي، يتم تشجيع الباحثين الذين يقومون بتقييم ثبات الاتساق الداخلي للاختبارات على استخدام البدائل، مثل أوميغا ماكدونالدز في المستقبل [31]. ورغم أننا لم نخض في الكثير من التفاصيل في دراستنا هذه حول الأمور المنهجية المتعلقة بالخصائص السيكومترية للاختبارات النفسية، إلاّ أننا لاحظنا أثناء محاولتنا الإجابة على تساؤلات هذه الدراسة؛ العديد من التجاوزات والأخطاء التي لم تكن في الأصل من أهداف دراستنا؛ ولكن لا يمكننا المرور دون التعقيب عليها، فقد وجدنا أن العديد من الخصائص السيكومترية للقياس، غالبًا ما لا يتم تقييمها أو تحليلها بشكل صحيح، ولا يتم الإبلاغ عنها بوضوح . وغالبًا ما تكون المعلومات المهمة حول صدق المحتوى مثلا محدودة للغاية . كذلك نادرا جدا ما يتم وصف الهدف من الاختبار، أي ما إذا كان مصممًا لأغراض تمييزية أو تقييميه أو غيرها، وهو الأمر نفسه الذي توصلت إليه دراسة [32] حيث أكدت أن هنالك العديد من التجاوزات التي تحدث أثناء التحقق من صدق وثبات الاختبار وجب ضبطها وتقييدها بمعايير موحدة وعالمية.

8 الاستنتاجات

نستنتج من خلال كل ما سبق ذكره ما يلي:

1. الاتساق الداخلي أقرب في معناه إلى الثبات أكثر منه إلى الصدق.
2. الاتساق الداخلي حتى وإن اعتبرناه جزافا صدقا، لا يمكن أبدا الاكتفاء به كدليل على صدق الاختبار بل وجب دائما مصاحبته بأدلة وبيّنات أخرى.
3. أكثر أدلة الصدق استخداما في الدراسات الأكاديمية العربية والأجنبية هي صدق الاتساق الداخلي وصدق المحكمين.
4. العديد من معاملات الارتباط المقبولة في الاتساق الداخلي ضعيفة جدا وأحيانا غير دالة.
5. استخدام الاتساق الداخلي كدليل صدق وثبات في آن واحد-باستخدام الاتساق الداخلي بين البنود أو مع المحور في الصدق واستخدام ألفا كرونباخ في الثبات- في العديد من الدراسات الأكاديمية.
6. ألفا كرونباخ أكثر أنواع الثبات استخداما في الدراسات الأكاديمية.

9 التوصيات والمقترحات

من خلال كل ما سبق التطرق إليه نوصي بالآتي:

1. التركيز على أن الصدق هو خاصية منطقية أكثر منها إحصائية عكس الثبات، وبالتالي يتوجب الاهتمام بالدلالة النظرية وعدم الاكتفاء بالدلالة الإحصائية.
2. لا يجب الاكتفاء بالاتساق الداخلي في دراسة الصدق، ولا يمكن استخدامه كإجراء للتحقق من الصدق والثبات في آن واحد.
3. هنالك العديد من الانتقادات التي وجهت لمعامل ألفا كرونباخ، ممّا يستوجب الحذر عند استخدامه.
4. عند استخدام الاختبارات النفسية سواء الجاهزة أو التي يتم بناءها من أجل الدراسة، وجب ذكر كل التفاصيل المتعلقة بها مثل المجتمع المستهدف، والغرض من الاختبار (تقييمي، تمييزي، تشخيصي وغيرها).

من ثم نقترح، الآتي:

1. ضرورة شد انتباه طلبة الدراسات العليا والباحثين إلى الأخطاء الشائعة في استخدامات الصدق والثبات.
2. البحث في أسباب شيوع استخدام معامل ألفا كرونباخ في الأبحاث الأكاديمية.
3. تعميم دراسة اشكالية الاتساق الداخلي (بين الصدق والثبات) على عينات كبيرة من الدراسات النفسية والاجتماعية (العربية والأجنبية).

References

[1] Cecil. *Reynolds and Ronald. Livingston, authored mastering modern psychometric theories and methods*, first edition. Amman : Dar Al-Fikr Publishers and Distributors, 2016-2017, (2013).

[2] Faraj. Safwat, *Psychometrics, Sixth Edition*. Egypt : The Anglo-Egyptian Library, (2017).

[3] Mohammed. Tighza, recent trends in estimating the validity and reliability of the measurement tools' scores : An applied evaluation theory analysis, *Journal of Psychological and Educational Sciences*. **4 (1)**, 7-29, (2017).

[4] Messick, S. Validity of Psychological Assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, **50 (1)**, 741-749, (1995).

[5] Sireci, S. On Validity Theory and Test Validation. *Educational Researcher*, **36 (8)**, 447-481, (2007).

[6] Anne Anastasi, and Susan Urbina. *Psychometrics,* 1nd ed. Arab Thought House Publishers and Distributors, Amman, Jordan, (2015).

[7] AREA, A. E., APA, A. P., & NCME, N. C. *Standards for Educational and Psychological Testing*. 3nd ed, Washington, American Educational Research Association, (2014).

[8] Cortina, J. What is coefficient alpha ? An examination of theory and applications. *J Apple Psychol*, **78 (1993)**, 98-104, (1993).

© 2022 NSP
Natural Sciences Publishing Cor.

[9] Bourque, P., & Beaudette, D. Etude psychometrique du questionnaire de depression de Beck aupres d'un echantillon d'etudiants universitaires francophones. Revue Canadienne des Sciences du Comportement **(14)**, 211-218, (2012).

[10] Furr, R., & Bacharach, V. *Psychometrics: An Introduction*, 2nd ed. Sage London, (2014).

[11] Fatima Abdul Rahim Al-Nawaisah, and Wejdan Khalil Al-Karaki. The level of psychological identity and life satisfaction among tenth grade students from their and their mothers' point of view : a comparative study. *Mutah for Research and Studies, Human and Social Sciences Series*, **35(2)**, 43-80, (2020).

[12] Rawan Abdrous Abdullah Al-Bar. *The effectiveness of a training program in early intervention based on verbal behavior in developing the communication skills of children with autism in the Kingdom of Saudi Arabia*. College of Education, United Arab Emirates University, UAE, (2016).

[13]Aisha Saif Sebaa Al Ali. *Differences in the cognitive and non-cognitive traits of female secondary school students according to the levels of mental and academic performance in the United Arab Emirates*. Bahrain: College of Graduate Studies, Arabian Gulf University, Bahrain, (2002).

[14] Rucker, R., Barlow, P. B., Bertolini, M., Carrera, E., Kossion, A., & Machini, L. Translation and preliminary validation of an ageism scale for dental students in brazil. *Gerodontology*, **37(1)**, 1-6, (2020).

[15] Abeer Mohammed Al-Sabban, and Samah Eid Al-Harbi. Students' addiction to using social networking sites and its relationship to psychological security and involvement in cybercrime. *International Journal of Educational and Psychological Studies*, **6(2)**, 267-293, (2019).

[16] Nawal Youssef Maatouk Abdullah. *Self-concept of mentally retarded and normal children and its relationship to behavioral problems.* College of Educational Studies, Arabian Gulf University, Bahrain, (1992).

[17] Weiner, D., Schneider, A., & Lyons, J. Evidence-based treatments for trauma among culturally diverse foster care youth: Treatment retention and outcomes. Children and Youth Services Review, **31(11)**, 1199-1205, (2009).

[18] Hadi Al-Huwailah. Satisfaction with life and its relationship to belief in self-efficacy, depression and psychological loneliness among a sample of people with visual impairment in the State of Kuwait. *Journal of Humanities and Social Sciences*, **15 (2)**,329-364, (2018).

[19] Whitfield, C. L., Dube, S. R., Felitti, V. J., & Anda, R. F. Adverse childhood experiences and hallucinations. *Child Abuse and Neglect*, **29(7)**, 797–810, (2005).

[20] Fang, Y., Niu, Y., & Dong, y. Exploring the relationship between narcissism and depression: The mediating roles of perceived social support and life satisfaction. *Personality and Individual Differences*, 173, (2021).

[21] Samar Issa Ibrahim Sabah. *The effect of a counseling program based on play and art in reducing aggressive behavior among children of SOS Children's Village in Bethlehem Governorate*. Faculty of Graduate Studies, Al-Quds Open University, Palestine, (2017).

[22] Mohamed Hussein Ahmed Nasser. School violence and its relationship to social skills among secondary school students in public schools in Ramallah and Al-Bireh. Faculty of Graduate Studies, Al-Quds Open University, Palestine, (2017).

[23] Abdul Rahman bin Safi, and Ali Tajin. Evaluation of methods for estimating the validity of the measurement tools used in psychological and educational research at the University of Algiers 2. *Suluk*, **3 (1)**, 32-59, (2013).

[24] Abdelaziz Bousalem. Cultural validity of psychological tests applied in Algeria and the problem of conditioning for validity. *Journal of Psychological and Educational Sciences*, **4 (2)**, 1-14, (2017).

[25] Wittkowski, A., Vatter, S., Muhinyi, A., Garrett, C., & Henderson, M. Measuring bonding or attachement in the parent-infant-relationship: A systematic review of parent-report Assessment mesures, their psychometric properties and Clinical utility. *Clinical Psychology Review*, **82 (1)**, 101-906, (2020).

[26] McAndrews, Z., Richardson, J., & Stopa, L. (2019). Psychometric properties of acceptance measures : A systematic review. *Journal of Contextual Behavioral Science*, **12(3)**, 261-277, (2019).

[27] Terwee, C. B., Schellingerhout, J. M., Verhagen, A. P., Koes, B. W., & de Vet, H. C. Methodological Quality of Studies on the Measurement Properties of Neck Pain and Disability Questionnaires: A Systematic Review. *Journal of Manipulative and Physiological*, **34 (4)**, 261-272, (2011).

[28] Irene, M., Luisa, S., Luca, G., & Alvisa, P. Instruments evaluating the quality of the clinical learning environment in nursing education: A systematic review of psychometric properties. *Inernational Journal of Nursing Studies*, **68 (4)**, 60-72, (2017).

© 2022 NSP
Natural Sciences Publishing Cor.

[29] Polit, D. Assessing measurement in health: Beyond reliability and validity, International. Journal of Nursing Studies, **52(1),** 1746-1753, (2015).

[30] Tian, L., Cao, X., & Feng, X. (2019). Evaluation of psychometric properties of needs assessment tools in cancer patients: A systematic literature review. PloS one, **14 (1)**, 210-242.

[31] Peters, G. The alpha and the omega of scale reliability and validity : Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*, **16 (2)**, 56-69, (2014).

[32] Terwee, C. B., Bot, S. D., de Boer, M. R., Van der windt, D. A., Knol, D. L., Dekker, J., et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology, **60 (1)**, 34-42, (2007).

© 2022 NSP
Natural Sciences Publishing Cor.