الصدق والثبات في القياس التربوي

1. الصدق:

يؤكد علماء النفس على أن خاصية الصدق من أهم خصائص الرائز الجيد وسماته، فبدونه لا يمكن الوثوق في قدرته على قياس ما صمم لقياسه.

وقد أشار أبو ناهية (1994) إلى أن الصدق هو الخاصية الأساسية الأولى التي يجب أن تتوفر في أداة القياس بصفة عامة، والروائز بصفة خاصة.

ومفهوم الصدق يشير إلى أن الرائز يقيس السمة أو الصفة التي صمم الرائز لأجل قياسها.

وقد أكدت النشرة الصادرة عن رابطة علم النفس الأمريكية (1995)، "APA" على ثلاث طرق أساسية تستخدم في حساب الصدق وهي:

- الصدق بدلالة المحتوى.
- الصدق بدلالة المحك.
- الصدق بدلالة التكوين.

1. صدق المحتوى: ويقصد به مدى تمثيل محتوى الاختبار للنطاق السلوكي الشامل للسمة المراد الاستدلال عليها، إذ يجب أن يكون المحتوى ممثلا تمثيلا جيدا للنطاق المفردات (المعارف، المهارات ...وغيرها) الذي يتم تحديده مسبقا.

حيث يعتبر هذا النوع من الصدق بمثابة حكم حول الدرجة التي يمكن لمحتوى الاختبار أن يتناسب مع السلوك المراد قياسه، وتشمل عملية تقدير صدق المحتوى على ثلاث خطوات:

أ. وصف مجال المحتوى.

ب. تحديد المجالات الفرعية التي تقيسها كل مفردة في الاختبار.

ت. مقارنة بنية الاختبار مع بنية مجال المحتوى، ورغم ما يبدو من بساطة هذه الخطوات إلا أنها تكون صعبة في تطبيقها، وأكبر صعوبة نواجهها تكون في الخطوة الأولى المتعلقة بوصف مجال المحتوى. وتوجد عدة طرق تجريبية للتحقق من صدق المحتوى وهي:

- المقارنة بين الدرجات التي يحصل عليها المفحوصين في الاختبار أو في صورة مكافئة له قبل التدريب وبعده للتحقق من التحسن الذي يطرأ على الاختبار نتيجة التدريب.
- دراسة أنماط الأخطاء الشائعة في الاختبار للتأكد من صدق مفتاح الاختبار.
- تحليل طرق العمل التي يستخدمها المفحوصون وذلك بإعطاء الاختيار فرديا مع توجيههم إلى التفكير بصوت مرتفع أثناء حل مشكلات الاختبار.

– دراسة أثر بعض العوامل المرتبطة بمحتوى ما يقيسه الاختبار

2. الصدق المرتبط بالمحك: ويحتوي هذا النوع من الصدق على:

أ .الصدق التلازمي:

هذا النوع من الصدق من أنواع الصدق التجريبي الذي يدل على وجود علاقة بين درجات الاختبار ومقاييس الأداء الراهن للفرد، والتي نحصل عليها في الوقت نفسه.

ب.الصدق التنبؤي:

ويُقصد به أنه قدرة الاختبار وفاعليته في التنبؤ بنتيجة معُيّنة في المستقبل.

3. صدق المفهوم:

ويشير صدق المفهوم بمدى نجاح الاختبار في قياس مفهوم فرضي معين من خلال تحقق من صدق المضمون البنود المكونة للاختبار ملائمة وموافقة للبناء(المفهوم) المقصود.

العوامل المؤثرة في صدق المقياس:

من أهم العوامل المؤثرة في صدق الاختبار ما يلي:

- **طول الاختبار:** إن عدد مفردات الاختبار تؤثر في صدقه، إذ أن العلاقة بينهما طردية، فكلما ازداد عدد البنود أو العبارات المستعملة فيه، ارتفع معامل صدق ذلك الاختبار، وذلك لأن أي ظاهرة لها مكوناتها الأساسية المتعددة تستوجب زيادة عدد البنود الاختبارية ذات العلاقة بتلك الابعاد، وهذا الأمر سيكون حتماً مؤثراً في طول إجراء هذه الاختبارات، والذي سيؤثر بدوره في صدقها.
- **معامل الثبات:** من بديهيات العمل الجماعي هو العلاقة ما بين صدق وثبات الاختبار، ومعامل الصدق يرتبط ارتباطاً مباشراً بثبات الاختبار، كما هو الارتباط بالمحك الذي يستعمل لحساب صدق الاختبار؛ ومما يجب التنويه إليه هو أن ثبات الاختبار أو المقياس يؤثر في صدقه، فانخفاض معامل الثبات دليل على وجود عيب في ذلك الاختبار، فليس كل معامل ثبات عالٍ لاختبارٍ ما سيكون دليلاً على نحو دائم على صدق ذلك الآخر.
- **عينة البحث:** حيث أن تطبيق الاختبار على عينة غير التي صمم ليتناسب مع مستواه وخصائصها يؤثر على صدق الاختبار ككل.
- **القصور في ضبط العوامل الدخيلة والمتغيرات التجريبية:** يحصل في بعض الأحيان أن يكون المعيار في الصدق التجريبي علاقة واضحة بالاختبار الذي وضعه الباحث، حيث أن عدم الدقة في ضبط العوامل الداخلية بين المعيار والاختبار يؤثر على معامل الصدق، ومثلا: الصعوبة أو السهولة المتناهية للاختبار أو عدم وجود رغبة لدى العينة في إجراء الاختبار أو ضعف محتويات الاختبار، كل ذلك يعد من العوامل الداخلية التي تؤثر كثيراً وتعمل على خفض معامل الصدق.

2. الثبات:

يعتبر الثبات الخاصية الثانية التي يجب أن يتصف بها الرائز الجيد بعد الصدق، لأن الصدق أكثر شمولية من الثبات، إلا أن هذا القول لا يعني الاستغناء عن الثبات إذا ما تحقق صدق الاختبار، وذلك لأننا لا نمتلك أدلة قاطعة على صدق الاختبار، لذا ينبغي تقدير الثبات فضلا عن تقدير الصدق.
والثبات يعني الاتساق في نتائج القياس على نفس الأداة أو الرائز عند تطبيقه من وقت لآخر.

كما عرفه عباس (1996) أنه: مدى تطابق درجات أفراد مجموعة معينة على رائز معين في كل مرة يعاد اختبارهم بنفس الرائز.

ومن أن الصفات الأساسية التي يجب توافرها في أداة جمع البيانات قبل الشروع في استخدامها هي خاصية الثبات، حيث تكمن أهمية قياس درجة ثبات أداة جمع البيانات في أهمية الحصول على نتائج صحيحة كلما تم استخدامه؛ فالأداة المتذبذبة لا يمكن الاعتماد عليها ولا الأخذ بنتائجها، وبالتالي ستكون نتائج الدراسة غير مطمئنة ومضللة، وفي أغلب الأحوال مضيعة للجهد والوقت والمال.
ومن طرق حساب ثبات الاختبار:

1. طريقة التطبيق وإعادة التطبيق Test -Retest:

تعتبر هذه الطريقة من أبسط الطرق وأسهلها في تعيين معامل ثبات الاختبار وتتلخص هذه الطريقة في تطبيق المقياس على مجموعة من الأفراد، ثم يعاد التطبيق مرة أخرى على نفس المجموعة، ويحسب معامل ارتباط بيرسون بين التطبيقين لتحصل على معامل ثبات درجات الاختبار، ويشير أحمد محمد الطبيب بأن معامل الثبات بإعادة التطبيق يطلق عليه اسم معامل الاستقرار، وهو يدل على استقرار عبر الزمن، ويتأثر معامل الثبات بعدد من العوامل أهمها أخطاء استقرار استجابات الأفراد وتصلح هذه الطريقة في حساب جميع أنواع المقاييس عدا اختبارات الذاكرة .

عيوبها: من عيوب هذه الطريقة صعوبة ضبط الظروف في التطبيقين وتأثر إجابات الأفراد خاصة في الاختبارات التحصيلية، بانتقال أثر التدريب وعاملي النضج والتعلم، كما أنها مكلفة في الوقت والجهد.

2. طريقة التجزئة النصفية:

في هذه الطريقة يتم تطبيق الاختبار على المفحوصين وبعد الانتهاء من التطبيق يتم تقسيم الاختبار إلى نصفين ثم يحسب معامل ارتباط بيرسون بين الدرجات على النصفين، بمعنى أننا نكون قسمنا الاختبار الى جزأين.
كما أنه وجب أن نراعي فيهما تساوي عدد الوحدات وكذلك القيم الإحصائية (المتوسط، الانحراف المعياري، معامل الارتباط) وهذا يستدعي ضرورة توفر شرط تجانس نصفي الاختبار. تتميز هذه الطريقة بتشابه ظروف التطبيق للأسئلة الفردية والزوجية وعدم التأثر بالممارسة والتدريب، وتوفر الوقت والجهد.

وعلينا ملاحظة أن معامل الارتباط المحسوب هو بين نصفي طول الاختبار، وهذه القيمة لا تنطبق مباشرة على الاختبار بكامله لذلك يتعين علينا تعديل هذا المعامل الناتج أو تصحيحه حتى نحصل على معامل ثبات الاختبار ككل، وهناك عدة معادلات تستخدم لتصحيح معامل ثبات نصفي الاختبار منها:

معادلة سبيرمان براون : Spearman Brown:

وفيها يتم التعويض بمعامل ارتباط بيرسون بين نصفي الاختبار لنحصل على معامل ثبات الاختبار ككل؛ وهي شائعة الاستخدام وبخاصة في اختبارات التحصيل والقدرات تحت ظروف محددة.

معادلة رولون: Rulon:

وتعتمد على حساب تباين درجات الاختبار ككل، ثم حساب تباين الفروق بين درجات الأفراد في النصف الأول، ودرجاتهم في النصف الثاني ثم تطبيق المعادلة نحصل على معامل ثبات الاختبار ككل.

معادلة جتمان Guttmann:

في هذه المعادلة يتم فيها حساب تباين درجات النصف الأول، وتباين درجات النصف الثاني، وتباين درجات الاختبار ككل. أي أنها تضع في الاعتبار احتمال اختلاف تباين درجات النصف الأول للاختبار عن تباين درجات النصف الثاني وهذا لا يتحقق في المعادلتين السابقتين.

3. طريقة الصور المتكافئة:

هذه الطريقة تتطلب إعداد من الاختبار الواحد صورتان متكافئتان (سواء كان بطارية أو مجموعة اختيارية معنية بقياس ظاهرة ما أو أي صفة تتعلق بتلك الظاهرة) من حيث تمثيل المتغيرات أو الوظائف التي تقاس. كما أن عدد المكونات الوظيفة في كل من الصورتين واحداً، بمعنى أن العناصر التي تقيس المكونات في الصورتين متماثلة، كذلك مستوى صعوبتها وسهولة وطريقة صياغتها، فضلاً عن تكافؤ الصورتين من حيث الطول وطريقة الإجراء والتصحيح والزمن المخصص للاختبار؛ حيث تطبق الصورتان على نفس المجموعة، ثم يحسب معامل الارتباط بين درجات كل من الاختبارين الممثلين لكلتا الصورتين، ويكون هذا المعامل هو المعبّر عن معامل الثبات.

4. طريقة الاتساق الداخلي:

تعتمد هذه الطريقة على استخدام كل من تباين فقرات الاختبار وتباين الاختبار ككل لتقدير الثبات وهي من طرق الأكثر شيوعا في تعيين الثبات؛ حيث تقوم على تقسيم الاختبار الواحد أكثر من مرة، وفي كل مرة بطريقة مختلفة وعند إجراء معامل الارتباط على كل نصفي الاختبار نحصل على تقدير مختلف للثبات باختلاف أساس التصنيف للاختبار كما يمكن أن يتم تجزئة الاختبار إلى عدد كبير من الأجزاء، حيث يتكون كل جزء من محور أو بعد واحد، ويراعى وجود تجانس داخلي بين المحاور أو الأبعاد، ويتحقق الثبات في هذه الطريقة من خلال عددا من المعادلات وضعها(كيودر- ريتشاردسون) لحساب الاتساق الداخلي بين كل أبعاد وعبارات الاختبارات وهي:

معادلة كيودر وريتشاردسون رقم (20) Kuder- Richarsdson:

وتعتمد على توفر البيانات عن تباين كل مفردة من مفردات الاختبار، وفي حالة عدم توفر هذه البيانات يمكن استخدام الصيغة رقم (21) بنفس المعادلة وهي تتميز بالسهولة والسرعة في حسابها.

حيث أنها تحتاج إلى معرفة تباين البنود ولكن من عيوبها أنها أقل دقة من الصيغ السابقة، وقد وضع **كيودر وريتشاردسون** شروطا لاستخدام هذه المعادلة وهي:

- أن تكون أسئلة الاختبار (صفر أو واحد).
- ألا يكون عدد الأسئلة المتروكة كبير.
- تقارب مستوى صعوبة الأسئلة.
- تساوي معاملات الارتباط بين درجات الأسئلة ومن الواضح أن هذه الشروط لا توجد في اختبارات مجال الصحة النفسية ولذلك ننصح بعدم استخدامها.

معادلة ألفا – كرومباخ Cronbach- Alpha:

يعتبر معامل ألفا كرومباخ حالة خاصة من معادلة كيودر وريتشاردسون ويتمثل معامل ألفا كرومباخ متوسط المعاملات الناتجة عن تجزئة الاختبار إلى طرق مختلفة، وبذلك فانه يمثل معامل الارتباط بين أي جزأين من أجزاء الاختبار ويتم حساب تباين كل بند من بنود الاختبار ثم مجموع التباينات، وكذلك الدرجة الكلية للاختبار، وتشترط أن تقيس بنود الاختبار سمة واحدة فقط، وتستخدم هذه المعادلة في المقاييس والاختبارات متعددة الاختبارات وليست الثنائية.

العوامل المؤثرة على الثبات:

هناك عدة عوامل تؤثر في معامل ثبات المقاييس منها:

- **طول الاختبار**: حيث تزداد قيمة معامل الثبات بزيادة عدد الأسئلة أو بنود الاختبار، ويرجع ذلك إلى أن زيادة عدد الأسئلة يؤدي إلى شمول أكثر للمحتوى ومن ثم صدق محتوى مرتفع، وهنا يمكن القول أن العلاقة بين عدد بنود الاختبار ومعامل ثباته علاقة طردية، بمعنى أنه إذا زاد عدد البنود ارتفع معامل ثبات الاختبار.
- **زمن الاختبار**: زيادة زمن الاختبار تؤدي بالفرد إلى الحصول على أعلى درجة متسقة مع قدرته، إلا أن زيادة الزمن بدرجة أكبر من اللازم قد تؤدي إلى الارتباك في الإجابة ومن ثم خفض معامل الثبات.
- **مستوى فقرات الاختبار**: الأسئلة الكثيرة الصعوبة أو السهولة على السواء كلاهما يقلل من ثبات الاختبار.