الوحدة الخامسة

الثبات والصدق

أولا: الثبات

يقصد بثبات الاختبار أو المقياس مدى إعطائه نفس الدرجات لنفس الأفراد عند تكرار تطبيقه عليهم؛ فالمقياس الثابت هو الذي إذا طبقته على فرد ثم أعدت تطبيقه على نفس الفرد بعد فترة معينة يعطيك تقريبا الدرجة التي أعطاها في المرة الأولى.

والواقع أنه ليس هناك اختبار أو مقياس يبلغ ثباته واحد صحيح؛ فالثبات نسبي ولكن ما نسعى إليه هو أن نضمن قدر الإمكان آلا يعكس الاختبار قدرا من التغير الطبيعي أكبر مما هو قائم بالفعل.

العوامل المؤثرة في ثبات المقياس

١. طول الاختبار

كلما كان المقياس أو الاختبار أكثر طولا؛ أي كلما زادت فقراته ارتفع معامل ثباته؛ فازدياد عدد الفقرات يعني أن الاختبار يقيس من السمة المراد قياسها قطاعا عريضا وأكثر اتساعا مما يقلل قدر الإمكان من تدخل عوامل المصادفة وغيرها من العوامل المؤقتة التي تؤثر على ثبات الاختبار، وعليه فإنه من الممكن أن نصل باختبار معين إلى درجة الثبات المطلوبة بإضافة عددا جديدا من الفقرات؛ ويمكن أن نعتمد في ذلك على المعادلة التالية:

مضاعفات الطول المطلوبة للاختبار =

$$\frac{\text{معامل الثبات المطلوب بلوغه (١ – معامل الثبات الحالي)}}{\text{معامل الثبات الحالي (١ – معامل الثبات المطلوب بلوغه)}}$$

مثال:

إذا كان هناك مقياس أو اختبار عدد فقراته (١٥ فقرة)، وكان معامل ثباته ٠,٤٠ ونريد الارتفاع بهذا المعامل ليصبح ٠,٨٠ نطبق المعادلة السابقة كالتالي:

$$\text{مضاعفات الطول المطلوبة} = \frac{٠,٨٠ (١ - ٠,٤٠)}{٠,٤٠ (١ - ٠,٨٠)} = \frac{٠,٤٨}{٠,٠٨} = ٦$$

أي علينا أن نضاعف فقرات المقياس ستة أضعاف لتصبح (٩٠ فقرة) حتى يصبح معامل الثبات ٠,٨٠.

٢. الزمن الذي يستغرقه المقياس أو الاختبار

يرتبط الزمن عادة بعدد الفقرات، وبما أن زيادة الفقرات ترفع ثبات المقياس أو الاختبار، فإنه بالتالي كلما زاد الزمن ازداد الثبات، على آلا يتجاوز الزمن الحد الأمثل.

٣. تشتت درجات الاختبار

إذ تؤدي زيادة الأسئلة شديدة الصعوبة والأسئلة شديدة السهولة إلى انخفاض معامل الثبات؛ في حين أن الأسئلة المتدرجة من حيث الصعوبة ترفع الثبات.

٤. وضوح صياغة الفقرات والتعليمات

فوضوح صياغة فقرات المقياس أو الاختبار بحيث لا تحتمل إلا معنى واحد يقصده الباحث، وكذلك التعليمات يجعل الأمر واضحا للمبحوثين جميعا مما يؤدي إلى زيادة الثبات.

٥. التخمين

فكلما زاد تأثر الفقرات بعامل التخمين نقص معامل الثبات نظرا لأن محددات التخمين بالنسبة للفرد في المرة الأولى تختلف عن المرة الثانية.

٦. حالة الفرد

تتأثر درجة ثبات المقياس أو الاختبار بكون الفرد في حالته الطبيعية ، والمقصود بالحالة الطبيعية أن يكون المبحوث وقت إجراء الاختبار في الحالة الصحية والنفسية التي أعتاد أن يكون عليها.

طرق قياس الثبات

أولا: طريقة الإعادة Test – Retest

تقوم هذه الطريقة على إجراء الاختبار على مجموعة من الأفراد، ثم إعادة تطبيقه مرة أخرى على نفس أفراد عينة التقنين Standardization Sample بعد انقضاء مدة مناسبة لا هي بالقصيرة فتسمح بتذكر الفرد لإستجاباته في المرة الأولى فيعيدها؛ فيبدو الاختبار ثابتا ثباتا مزيفا، ولا هي بالطويلة فتسمح بحدوث تغيرات جوهرية في استعدادات الفرد أو الخاصية المقاسة، فيبدو الاختبار ضعيف الثبات حتى لو كان ثباته مرتفعا أصلا.

ويحدد الفترة المناسبة بين التطبيقين تقدير مؤلف أو معد الاختبار أو المقياس وفقا لطبيعة المقياس والعينة وظروف التطبيق، على آلا تنقص هذه الفترة عن أسابيعا قليلة بالنسبة للصغار، وآلا تتجاوز ستة أشهر بالنسبة للكبار.

ويستخرج معامل الثبات في هذه الطريقة بحساب معامل الارتباط Correlation Coefficient بين درجات التطبيق الأول والثاني لنفس الأفراد، ويكون معامل الارتباط هنا هو

نفسه معامل ثبات الاختبار أو المقياس، وكلما كان المعامل موجبا مقتربا من الواحد الصحيح كلما دل على ارتفاع ثبات المقياس أو الاختبار، وكلما كان منخفضا أو مقتربا من الصفر أو سالبا كلما دل على ضعف ثبات الاختبار.

وعادة يعد الثبات مرضيا لو بلغ (+٠,٧٠) فأكثر، ويعد ثبات الإعادة هو أفضل أنواع الثبات لأنه بالفعل يجسد مضمون الثبات ويعبر عنه.

ثانيا: طريقة التقسيم النصفي Split – Half

تقوم هذه الطريقة على تطبيق المقياس أو الاختبار على عينة التقنين مرة واحدة ثم يحسب معامل الارتباط بين إستجابات الأفراد على الأسئلة الزوجية وإستجاباتهم على الأسئلة الفردية، وميزة هذه الطريقة أنها تستبعد العيوب التي يمكن أن تنجم عن طريقة إعادة الاختبار.

وتشترط هذه الطريقة أن تتساوى القيم العددية للمقاييس الإحصائية المختلفة لنصفي الاختبار وذلك يعني:

١. تكافؤ الجزئين من حيث مستويات صعوبة الأسئلة؛ بمعنى أن صعوبة السؤال الأول في النصف الفردي تساوي صعوبة السؤال الأول في النصف الزوجي.
٢. تساوي متوسط درجات الأفراد على النصفين.
٣. تساوي تشتت درجات الأفراد على النصفين.

والواقع أن تجزئة الاختبار على هذا النحو (فردي وزوجي) تعد الأصح؛ فنحن نعلم أنه لا يمكننا قسمة المقياس إلى قسمين وفقا لترتيب الأسئلة نظرا لعدم تجانسها، فعادة ما يبدأ المقياس أو الاختبار بالأسهل وينتهي بالأصعب، هذا فضلا عن تدخل تأثيرات الحماس والتعب والملل وغير ذلك من العوامل التي تختلف في بداية الاختبار عنها في نهايته، لذلك فالتقسيم على أساس زوجي وفردي يعد بمثابة الأسلوب الأمثل، والاعتراض الرئيس الذي يوجه إلى هذه الطريقة يتمثل في كونها لا تتيح لنا سوى الحصول على معامل ثبات نصف الاختبار فقط، ولأن معامل الثبات يتأثر بعدد الفقرات أصبح حتما تحويل معامل الثبات من معامل نصفي إلى معامل ثبات الاختبار ككل، وثمة معادلات تستخدم لهذا الغرض من أهمها معادلة (سبيرمان – براون) Spearman – Brown ، والتي تحسب وفقا للقانون التالي:

$$\text{معامل الثبات بعد التصحيح} = \frac{\text{معامل الثبات النصفي} \times ٢}{\text{معامل الثبات الكلي} + ١}$$

مثال:

إذا تم حساب معامل الثبات بطريقة التنصيف فكان ٠,٦٠ مثلا، فإن تحويله إلى معامل ثبات الاختبار ككل يكون كالتالي:

$$\text{ر} = \frac{٢ \times ٠,٦٠}{١ + ٠,٦٠} = ٠,٧٥$$

ثالثا: طريقة الصور المتكافئة Equal Form

في بعض الحالات لا يكتفي معد الاختبار أو المقياس بعمل صورة واحدة من المقياس، بل يعد أكثر من صورة بديلة بحيث إذا تعذر تطبيق أحدهما حلت الأخرى محلها بنفس الكفاءة، والصورة المتكافئة للاختبار أو المقياس ينبغي أن تتعادل معه في كافة الخصائص الأساسية مثل (شكل الفقرات – عددها – مستوى صعوبتها – طريقة التطبيق – طريقة التصحيح .. إلخ)، وفي هذه الحالة يتم تطبيق كلا الصورتين في نفس اليوم على نفس أفراد عينة التقنين؛ ثم نحسب معامل الارتباط بين الدرجتين فيكون هذا هو معامل الثبات، وتمتاز هذه الطريقة في دراسة الثبات بأنها تلغي تأثير الذاكرة في طريقة الإعادة كما تجنب الباحث الصعوبات والمزالق التي يتعرض لها في تحديد مدى طول الفترة المنقضية بين التطبيق الأول والإعادة.

رابعا: طريقة كودر – ريتشاردسون Kuder – Richardson

ويطلق عليها طريقة تحليل التباين؛ إذ أن هذه الطريقة تعتمد على مدى الاتفاق بين الإستجابات على كل وحدة في الاختبار أو المقياس وغيرها من الوحدات، أي تشير إلى مدى الثبات الداخلي للمقياس.

ويحسب معامل الثبات هنا بالتعويض في المعادلة التالية:

$$\frac{\text{عدد البنود} \times \text{تباين درجات الاختبار} - \text{المتوسط (عدد البنود - المتوسط)}}{(\text{عدد البنود} - ١)\ \text{تباين درجات الاختبار}}$$

حيث أن :

- تباين درجات الاختبار = ١٠٠/١ (عدد الأفراد × مجموع المربعات – مربع الدرجات)
- المتوسط = $\frac{\text{مجـ س}}{\text{ن}}$

مثال:

فيما يلي درجات حصل عليها أفراد في اختبار يتكون من ٢٠ بندا:

الأفراد	الدرجات	المربعات
١	١١	١٢١
٢	٥	٢٥
٣	١٢	١٤٤
٤	٧	٤٩
٥	١١	١٢١
٦	١٢	١٤٤
٧	١٦	٢٥٦
٨	٧	٤٩
٩	١٦	٢٥٦
١٠	٥	٢٥
مجموع	١٠٢	-
مربع الدرجات	١٠٤٠٤	
مجموع المربعات		١١٩٠

تباين درجات الاختبار = ١/١٠٠ (١٠ × ١١٩٠ - ١٠٤٠٤) = ١٤,٩٦

المتوسط = $\frac{١٠٢}{١٠}$ = ١٠,٢

معامل الثبات =

$$\frac{٢٠ \times ١٤,٩٦ - ١٠,٢ (٢٠ - ١٠,٢)}{(٢٠ - ١) ١٤,٩٦} = \frac{٢٩٩,٢ - ٩٩,٩٦}{٢٨٤,٢٤}$$

معامل الثبات = $\frac{١٩٩,٢٤}{٢٨٤,٢٤}$ = ٠,٧٠

Dr. Mohsen Lotfy Ahmed

ثانيا: الصدق

يقصد بصدق الاختبار أو المقياس مدى قياسه فعلا للاستعداد أو الخاصية التي وضع لقياسها؛ أو بمعنى أخر ، أن الصدق تقدير لمعرفة ما إذا كان الاختبار يقيس ما نريد أن نقيسه به، وكل ما نريد أن نقيسه به، ولا شيئ غير ما نريد ان نقيسه به أم لا ؟.

وتتعدد أساليب حساب وتقدير الصدق في ضوء هذا التعريف العام، فنحصل في بعض الحالات على معامل كمي للصدق بالتعبيرات الرقمية أو نحصل في حالات أخرى على تقدير كيفي للصدق بالتعبيرات اللفظية.

ومن أساليب حساب الصدق:

١. الصدق السطحي أو المظهري Face Validity

يقصد بالصدق السطحي صدق المظهر العام للمقياس أو الاختبار أو مدى العلاقة بين بنوده والمتغير الذي يقاس، إذ أنه لا يزيد عن النظرة التأملية النظرية الصرفة للمقياس لتخمين مدى علاقة محتواه بالاستعداد أو الخاصية التي وضع لقياسها.

ومن المرغوب فيه أن يكون المقياس ذا صدق مظهري إذ يلعب هذا الصدق دورا في تنمية تعاون المبحوث وفي توجيه انتباهه إلى نوعية الإجابة المطلوبة، إلا أن الصدق السطحي لا يعد وحده محكا للصدق، فهو ليس صدقا بالمعنى العلمي للكلمة، ففضلا عن اعتماده على التقدير الذاتي للقائمين على الاختبار؛ فهو من الناحية الموضوعية ينصب على ما يبدو أن الاختبار يقيسه وليس على ما يقيسه بالفعل.

٢. صدق الارتباط بمحك Criterion Related

ويعد من أفضل أنواع دراسة الصدق؛ فالمحكات التي يمكن أن تمثل الاستعداد أو الخاصية في الواقع الخارجي تعد أفضل ميزان لتقدير مدى صدق الاختبار، ونحصل على معامل الصدق في هذه الطريقة عن طريق إيجاد معامل الارتباط بين درجات عينة التقنين على المحك ودرجاتهم على الاختبار، ويصبح معامل الارتباط هنا هو معامل الصدق.

ومن أهم المحكات التي تصلح لدراسة صدق الاختبار أو المقياس:

أ. مستوى الأداء في العمل
ب. التحصيل الدراسي
ج. المجموعات المتناقضة
د. التفرقة بين الأعمار
ه. التقدير الشخصي
و. الارتباط بين الاختبار واختبارات أخرى صادقة

Dr. Ihsan Al-Hamdan

٣. الصدق العاملي Factorial Validity

يعني الصدق العاملي لاختبار ما تشبع Saturation درجة هذا الاختبار بالعامل Factor المفترض أنه وضع لقياسه، وكلما كان التشبع كبيرا دل ذلك على ارتفاع مستوى صدق الاختبار.

ولحساب الصدق العاملي لاختبار نختار عدة اختبارات من الاختبارات الصادقة والصالحة لقياس الاستعداد الذي وضع الاختبار لقياسه ثم نقوم بتطبيقها مع الاختبار الجديد سويا على كل فرد من أفراد عينة التقنين ونحسب معاملات الارتباط بين درجات كل اختبار والأخر ثم نجري على هذه المصفوفة الارتباطية الطريقة الإحصائية المعروفة بالتحليل العاملي لنعرف منها مدى تشبع الاختبار الجديد بالعامل العام الشائع بين هذه الاختبارات.

المراجع:

- الغريب، رمزية ١٩٨٥ ، التقويم والقياس النفسي والتربوي، القاهرة، دار النهضة العربية.
- طه، فرج عبد القادر٢٠٠٦، أصول علم النفس الحديث، الرياض ، دار الزهراء للنشر والتوزيع.
- فرج، صفوت أرنست ١٩٨٩، القياس النفسي، القاهرة ، مكتبة الأنجلو المصرية.
- لطفي، محسن ٢٠١١، مقدمة في الإحصاء الاجتماعي، الرياض، المجلس العلمي للطباعة والنشر، جامعة الملك سعود.
- Anastasi, Anna 1990, Psychological testing, New York., Macmillan publishing.
- Gregory, R. J.1996, Psychological testing, History principles and Applications, Allyn & Bacon, U.S.A.

Dr. Mohsen Lotfy Ahmed