المحاضرة الخامسة (5) --- صفات الاختبار الجيد ...(الثبات)

ثالثا - ثبات الاختبار Test Reliability:

بحسب أنستازي (1988، ص. 109) يشير الثبات إلى اتساق الدرجات المستخرجة من الأشخاص أنفسهم عندما يعاد اختبارهم بالاختبار ذاته في مناسبات مختلفة، أو عندما يختبرون بمجموعات مختلفة من بنود متكافئة، أو حينما يختبرون في ظل متغيرات لأخرى (عبد الخالق، 2000. ص. 123).

يتصف الاختبار بالثبات عندما يعطي نفس النتائج أو نتائج متقاربة إذا طبق أكثر من مرة في ظروف مماثلة. فلو استخدم اختبار قياس ذكاء طالب ما وحصل على درجه 120 فأن هذا الطالب يجب أن يحصل على نفس النتيجة تقريباً لو تقدم لنفس الاختبار بعد أسبوعين أو شهر مثلاً.

⋆ طرق حساب الثبات:

أ- ثبات إعادة الاختبار Test- retest:

يدل على الاستقرار عبر الزمن، ويتلخص في تطبيق الاختبار أو المقياس على الأفراد ذاتهم، ثم يتركون مدة من الزمن، ويعاد تطبيق المقياس على الأفراد أنفسهم ويحسب معامل ارتباط بين التطبيقين وهو معامل الاستقرار stability coefficient. والنقطة المهمة في هذه الطريقة هي تحديد الفاصل الزمني الأمثل بين التطبيقين، بحيث لا يكون قصيرا جدا حتى يخشى معه انتقال آثار التعلم أو التدريب أو التذكر. ومن ناحية أخرى لا يكون طويلا جدا إلى الدرجة التي يخشى معها تغير في السمة موضع القياس، فنكون في هذه الحالة بصدد قياس ثبات لكل من الأداة والسمة معا (عبد الخالق، 2000، ص. 125).

تتطلب هذه الطريقة اعادة تطبيق الاختبار مرة أخرى على أفراد المجموعة نفسها بعد فترة زمنية ملائمة، ثم تحسب بعد ذلك معامل الارتباط بين الدرجات التي حصل عليها أفراد العينة في المرة الأولى والثانية، ويسمى معامل الارتباط المستخرج بهذه الطريقة بمعامل الاستقرار، أي استقرار نتائج الاختبار خلال الفترة بين التطبيقين للاختبار (الريماوي، 2017، ص. 127).

يوجد عدد من العيوب وجهت لهذه الطريقة وهي بأننا لا نضمن أن تكون ظروف إجراء الاختبار الأول هي نفس ظروف تطبيق الاختبار الثاني، كما أنه قد يؤدي إلى ألفة المختبرين للاختبار في المرة الثانية، بالإضافة لكثرة تكاليفها وطول الوقت المستخدم (الكبيسي، وربيع، 2008).

ب- ثبات التجزئة النصفية split- half:

من الممكن بعد تطبيق واحد لصيغة واحدة من الاختبار أن نصل إلى اختبار أو مقياس للثبات عن طريق قسمة الاختبار إلى نصفيين، حيث تستخرج درجتان لكل شخص بقسمة أدائه على الاختبار نصفين متساويين. ويمدنا ثبات القسمة أو التجزئة النصفية بمعامل لاتساق عينات محتوى الاختبار، ولا يدخل استقرار الدرجات عبر الزمن في هذا النوع من الثبات، لأن جلسة القياس واحدة فقط، ويسمى هذا النوع من معامل الثبات أحيانا بمعامل الاتساق الداخلي internal consistency لأنه يتطلب تطبيقا واحدا فقط لصيغة واحدة من الاختبار (عبد الخالق، 2000، ص. 127).

لا تحتاج هذه الطريقة إلى تطبيق الاختبار مرتين، ولا إلى تصميم صورتين متكافئتين للاختبار، وإنما تتطلب تصميم اختبار متكافئين، على أن يحتوي القسم الأول على الدرجات الفردية 1. 3. 5. 7 والقسم الثاني على الدرجات الزوجية 2. 4. 6. 8 وتمتاز هذه الطريقة بأنها تتغلب على العوامل المؤثرة في أداء الأفراد مثل الوقت والجهد والتعب، كما أنها تمتاز بتوحد ظروف الإجراء توحيدا كاملا. ومن الشروط الواجب اتباعها في تصميم الاختبار عند استخدام هذه الطريقة أن تتكافأ البنود الفردية مع البنود الزوجية. وللحصول على الثبات نستخدم معاملات الارتباط بين درجات كل فرد على البنود الفردية والزوجية، ويستخدم بعد ذلك معادلة سبيرمان براون (الريماوي، 2017، ص. 129).

وطريقة التجزئة النصفية لمعامل الثبات تتساوى نظريا مع طريقة الصور المتكافئة ومع ك تعتبر قياسا للاتساق الداخلي لأن الصورتين المتساويتين متضمنتان في اختبار مفرد؛ بمعنى أننا بدلا من تطبيق الصورتين للاختبار بالتناوب فإننا نكتفي بتطبيق اختبار واحد في احتساب معامل الثبات عن

طريق الحصول على الدرجات البديلة لكل من النصفين ومن ثم احتساب معامل الارتباط لدرجات هذين النصفين، وفي معظم الحالات يستخدم معامل ارتباط بيرسون للحصول على معامل الثبات في هذه الحالة أما إذا أردنا الحصول على الثبات للاختبار كله (وليس لنصفه) ففي هذه الحالة يستحسن استخدام معادلة سبيرمان– براون (الصراف، 2002).

ج- ثبات كودر- رتشاردسون Kuder - Richardson ومعامل ألفا Cronbach:

يتلخص في تطبق واحد لصيغة واحدة للاختبار، وبيان مدى الاتساق في الاستجابات لكل بنود الاختبار، أي التأكد من قياس كل الأجزاء المكونة للاختبار للشيء نفسه، ولذلك يعطي درجة (اتساق ما بين البنود) inter- item consistency بعد فحص الأداء على كل بند. وقد وضع كرونباخ معادلة أعم تصلح لاستخبارات الشخصية تدعى معامل ألفا (عبد الخالق، 2000، ص. 128).

د- طريقة الصور المتكافئة Parallel Forms:

تعد هذه الطريقة من أفضل الطرق ملائمة مع الاختبارات التحصيلية ويفترض تكوين صورتين متكافئتين من الاختبار الواحد بحيث تكون هذه الصور متكافئة، وتتوفر فيها نفس مواصفات الاختبار الذي نريد التحقق من ثباته بحيث يحتوي على نفس العدد من الأسئلة وأن تكون صياغة الأسئلة متماثلة ودرجة الصعوبة واحدة، وأن تتضمن محتوى واحد، وأن تتفق معه في جميع المظاهر الأخرى مثل التعليمات والأمثلة والشكل العام (الريماوي، 2017، ص. 125).

والتكافؤ يعني إعداد اختبارين لهما نفس الخصائص من حيث صعوبة السؤال وقدرته التمييزية وتشابه المحتوى الذي يقيسه السؤال وتشابه في انسجام السؤال مع الهدف الذي يقيسه، هذا على مستوى كل سؤال من الفقرات في الاختبارين وهناك شروط للتكافؤ على المستوى الإجمالي للاختبار، وهي أن يتشابه شكل التوزيع للعلامات على الاختبارين وأن يكون لهما نفس الوسط الحسابي ونفس التباين، وليس المقصود التطابق في القيم ولكن المقصود أن يكون الفرق غير دال إحصائيا (عودة، 2000).

⋆ عوامل تؤثر على ثبات الاختبار:

يشير كل من (الشيخ وآخرون، 2004)، و(النبهان، 2004)، و(علام، 2000)، و(أبو لبدة، 1996)، و(إسماعيل، 2007)، و(عوض، 1999) إلى أهم العوامل التي تؤثر في ثبات الاختبار، وهي:

أ- طول الاختبار: يزداد ثبات الاختبار بزيادة طول الاختبار ويقل ثباته بالعكس ولكن بشرط ألا يؤدي طول الاختبار إلى إثارة الملل عند المفحوصين.

ب- زمن الاختبار: يزداد ثبات الاختبار بزيادة الوقت الذي يستغرقه المفحوص في أداء الاختبار ويقل الثبات بانخفاض مدة الاختبار.

جـ- تجانس المفحوصين: يزداد ثبات الاختبار إذا كأن المفحوصين أقل تجانساً ومن مستويات مختلفة، إما إذا كأن المفحوصين متجانسين ومتقاربين في المستوى من السلوك الذي يقيسه الاختبار فأن درجة الثبات ستقل وذلك لأن المفحوصين المتجانسين يحصلون على درجات متقاربة يمكن أن يتغير ترتيبها عند إعادة تطبيق الاختبار.

د- مستوى صعوبة الاختبار: يزداد ثبات الاختبار كلما زادت سهولته لأن ذلك يفقده القدرة على التمييز، كما يقل الثبات إذا ازدادت صعوبة الاختبار لأن ذلك سيدفع المفحوصين إلى التخمين. فالأسئلة السهلة والصعوبة عادة يأخذ المفحوص عليها علامات متقاربة.

هـ- التمييز: الاختبار المميز هو الاختبار الذي يستطيع أن يبرز الفروق بين المفحوصين أو بين التلاميذ المتفوقين والضعاف لذلك ينبغي أن تكون جميع الأسئلة التي يتضمنها الاختبار مميزة أي أن كل سؤال أو فقرة تختلف الإجابة عليه باختلاف التلاميذ وهذا يتطلب أن يكون هناك مدى واسع بين الفقرة السهلة والفقرة الصعبة بحيث يؤدي هذا إلى توزيع معتدل بين أعلى وأقل درجات وأن تصاغ الأسئلة في كل مستوى من مستويات الصعوبة بحيث يحصل التلاميذ على درجات متفاوتة.

و- التخمين: يلجأ بعض من المفحوصين في حالة عدم تأكدهم من الإجابة الصحيحة إلى التخمين، مما يؤدي إلى خفض ثبات الاختبار، حيث زيادة أثر التخمين يقود إلى نقص الثبات.

ز- موضوعية الاختبار: تعد موضوعية التصحيح من بين العوامل المؤثرة في ثبات لاختبار، وعلى الأخص في الاختبارات التي تعتمد على تقدير المصحح كاختبارات المقال التحصيلية واختبارات الإبداع، حيث أن تباين التصحيح يؤثر في زيادة تباين الخطأ، وبالتالي إلى نقصان معامل الثبات.

فعندما تكون الأسئلة غامضة وطويلة تقلل من الثبات، والأسئلة الواضحة والموضوعية والقصيرة تزيد من الثبات، وهذا يتطلب من الباحث الدقة في اختيار الألفاظ والعبارات ونوعها، عند صياغة الأسئلة حتى يجعلها قادرة على الوصول إلى الثبات الحقيقي.