الفصل الثانى

الثبات

١- مقدمة

٢ - نظرية الثبات

٣ - مفهوم الثبات النسبى وطرق حسابه

- طريقة إعادة الاختبار
- طريقة الصور المتكافئة
- طريقة التجزئة النصفية
- طريقة الاتساق الداخلى
- طريقة تحليل التباين

٤ - مفهوم الثبات المطلق وطريق حسابه

٥ - العوامل التى تؤثر على ثبات درجات الاختبار

أولا : العوامل الخارجية

ثانياً : العوامل الداخلية

٦- التصحيح للتنقية

٧ - علاقة الموضوعية بالثبات

- معجم إنجليزى - عربى للمصطلحات
- مراجع الفصل الثانى

obeikandl.com

obeikandl.com

١- مقدمـــة

يواجه القـائميين بالقيـاس فى مجال التربيـة البدنية والرياضة الكـثير من المشكلات التى تُعـرِّض عملية القياس لـلعديد من الصعوبات التـى قد يترتب عليها حدوث أخطاء فى عملية القياس. فقـد تستهدف عمليات القياس تقدير خصائص أو صفات بسيطة واضحة ومحددة مثل قياس بعض أبعاد الجسم، أو تستهـدف قياس بعض الصفات أو الـسمات أو القدرات المعقـدة مثل الذكاء، والقلق، والعدوان، واللياقة البدنية، أو القدرة الحركية وغيرها.

ومن الملاحظ أن القياس فى التـربيـة البدنية والرياضيـة قليـلا ما يتعامل مع الصفـات أو الخصائص البـسيطة Simple Qualities، فهو يتـعامل - فى معظم الحـالات - مـع العديـد مـن الصفـات والخصائـص والسمـات المعقـدة Complex traits، وهى لسوء الحظ صفات وخصائص وسمات لا ترى بالعين المجردة ولايـمكن حسـها أو لمسـها باليـدين، وإنما هى صـفات وخصـائص وسمـات يمكن التـعبـير عنهـا من خلال مظـاهر السلوك (الأداء) الـتى تدل عليـها، فى ذات الوقت فإنـه لا يمكن استـخدام المسطرة المدرجـة أو شريط القياس لقياس مثل تلك الصفات والخصائص.

وقد اهتم علماء القـياس بتقويم الأدوات التى يستخدمـونها سواء أكانت هذه الأدوات المسطرة المدرجة أو شريط القيـاس أو الميزان الطبى أو غيرها من أدوات ووسائل القياس، وفى هـذا الصدد يوضح نانلى Nunnally ١٩٧٨م أن أدوات القياس سواء استخدمت لقياس صفات بسيطة أو معقدة يجب أن تكون قادرة على إعـطاء نتائج دقيقـة فى كل مرة تستـخدم فيها مـا دامت الشروط موحدة فـى جميع مرات الـتطبيق. فإذا كـان شريط القيـاس المستخـدم قابلا للاستطالة (المط) أثناء التطبيق، فإنـه سوف يعطينا إما قراءات أكثر أو أقل من القراءات الحقيقية للظاهرة المقيسة.

من هذا المنطلق يحاول علماء القياس فهم سلوك الإنسان Human behavior على أساس تقويم ثبات الكثير من الاختبارات والمقاييس حتى لا يقعوا فى الأخطاء التى يمكن أن يتعرض لها «النجار» الذى يحاول بناء منزل من الخشب باستخدام شريط قياس من المطاط لن يعطيه أبداً أطوالا موحدة لنفس قطع الخشب المطلوب أن تكون متساوية فى الطول.

٢ - نظرية الثبات

يمكن القول بأن نظرية الثبات Reliability theory تعنى نظرية خطأ القياس the theory of measurement error، وهى نظرية تم تطويرها خلال المحاولات التى تمت فى مجال القياس النفسى psychometrics، والتى كانت تستهدف دائماً مراجعة وتقويم الثبات حتى تكون المقاييس measurements المتاحة للاستخدام يعتد بها من حيث الثقة فى النتائج التى يتم الحصول عليها.

وأخطاء القياس measurement errors ليست قاصرة على القياس التربوى والنفسى، ولكنها تحدث بالنسبة للقياس فى مجال علوم التربية البدنية والرياضة، وفى مجال العلوم الاجتماعية والبيولوجية وغيرها، فمن المعروف أن مقاييس measures الناتج القومى National product فى مجال الاقتصاد، ومقاييس ضغط الدم blood pressure فى مجال الطب تتعرض لأخطاء القياس كما يحدث بالنسبة للاختبارات التربوية والنفسية برغم الاهتمام الشديد بإعداد مثل هذه الاختبارات. وبسبب ظاهرة أخطاء القياس سوف يظل الثبات يستحوذ على اهتمام علماء القياس فى المجالات المختلفة.

دراسات سبيرمان المبكرة:

لقد حقق علم النفس تقدماً مشهوداً بالنسبة لتطوير ثبات أدواته ووسائله المختلفة وكان الفضل فى ذلك يرجع إلى المحاولات المبكرة التى قام بها عالم

النفس الإنجليزى الشهير تشارلز سبيرمان C. Spearman والذى استفاد فيها بما أشار إليه سلفه دى موفر De movire عن خطأ المعاينة Sampling error وهو خطأ ينجم عن اختيار عينة لا تمثل مجتمعها الأصلى. كما استفاد سبيرمان كذلك من الجهود التى قدمها كارل بيرسون Karl pearson فى إعداده لمعامل الارتباط التتابعى product moment correlation والذى نشره عام ١٨٩٦م. فقد وظف سبيرمان هذين الانجازين معا فى سياق عملية القياس ليضع بذلك معظم الأسس التى نعرفها الآن بالنسبة لنظرية الثبات والتى نشرها لأول مرة فى عام ١٩٠٤م تحت عنوان:

"The proof and measurement of association between two things"

ولأن مجلة علم النفس الإنجليزية British Journal of psychology لم تبدأ فى الصدور إلا عام ١٩٠٧م، لذا فقد قام سبيرمان بنشر أعماله التى توصل إليها فى مجلة علم النفس الأمريكية American Journal of psychology عام ١٩٠٤م.

وعندما نشر سبيرمان مقالته فى مجلة علم النفس الأمريكية فى يناير عام ١٩٠٤م لفتت انتباه رائد القياس النفسى فى أمريكا ادوارد ل. ثورنديك Edward L. Thorndike الذى كان يعد الطبعة الأولى من كتابه الشهير «مقدمة فى نظرية المقاييس العقلية والاجتماعية "Introduction to the theory of mental and social measurements".

وقد كان كتاب ثورنديك من الأعمال الأولى الرائدة التى ظهرت فى بداية القرن العشرين والتى تميزت بمستوى علمى رفيع لفت أنظار كل معاصريه من علماء القياس، فقد كان من أهم نتائج هذا الكتاب هو بدء الاهتمام بموضوع الثبات فى مجال القياس ليس فى أمريكا فحسب ولكن فى أوروبا أيضاً.

وبالإضافة إلى إنجازات كل من سبيرمان وثورنديك بالنسبة لنظرية الثبات تأتى إسهامات كل من كيودر، وريتشاردسون Kuder and Richardson ١٩٣٧م. وذلك بما قدماه من بعض الطرق الجديدة لتقدير معاملات الثبات

Reliabitity coefficients، والتى انتهت بما قدمه كرونباخ Cronbach ١٩٧٢م ومعاونوه من تطوير طرق تقويم العديد من مصادر الخطأ Sources of error فى البحوث السلوكية.

مفهوم الثبات:

ثبات القياس Reliability of measurement يعنى أن الدرجات التى يتم الحصول عليها دقيقة accurate وخالية من الخطأ error، وهذا يعنى أنه فى حالة تطبيق نفس أداة القياس (الاختبار أو المقياس) على نفس الفرد أو الشىء أى عدد من المرات بنفس الطريقة والشروط، فإننا سوف نحصل على نفس القيمة فى كل مرة، حيث تدل هذه القيمة على أن الشىء الذى تم قياسه لم يتغير فى غضون فترات القياس المختلفة.

ومن المتفق عليه أن الثبات يشير الى مدى الدقة والإتقان أو الاتساق الذى يقيس به الاختبار الظاهرة التى وضع من أجلها، لذلك يمثل الثبات أحد أهم الشروط الواجب توافرها فى أدوات القياس أيا كانت هذه الأدوات، وذلك لكونه - أى الثبات - يشير إلى مدى الوثوق فى الدرجات التى يتم الحصول عليها من تطبيق الأداة، بمعنى عدم تأثير الدرجات التى يتم الحصول عليها من تطبيق الأداة بأخطاء الصدفة، حيث يطلق على الثقة فى نتائج الأداة اسم: ثبات درجات الاختبار أو المقياس، فنحن لا نستطيع أن نثق فى اختبار للعدو طبق على تلميذ فى المرحلة الإعدادية فكان زمنه ٤ ثوان، وبعد أسبوع طبق نفس الاختبار على نفس التلميذ فكان زمنه ٧ ثوان. فكأن الثبات بهذا المعنى يدلل على الضمانات الخاصة بالحصول على نفس النتائج - تقريبا - إذا أعيد تطبيق الاختبار أو المقياس على نفس مجموعة الأفراد مع توحيد كافة الشروط والمتغيرات فى مرات التطبيق المختلفة، وهو ما يؤكد ضعف تأثير عوامل الصدفة أو العوامل الدخيلة أو العشوائية على نتائج الاختبار أو أداة القياس.

ومهما تعددت مفاهيم الثبات، فإنه يمكن إيضاح معناه فى النقاط التالية:

١- يختص الثبات بالنتائج التى يتم الحصول عليها من الاختبار أو المقياس، ولا يختص بالاختبار أو المقياس نفسه، وبناء على ذلك يصبح من الأكثر دقة أن نشير إلى ثبات درجات الاختبار أو المقياس بدلا من الإشارة إلى ثبات الاختبار أو أداة القياس.

٢- يبين الثبات مدى تأثير عوامل الصدفة (العشوائية) على نتائج الاختبار أو المقياس، وعليه فكلما كان الثبات مرتفعًا كان ذلك تأكيدًا على ضعف تأثير عوامل الصدفة، وكانت الدرجات التى يتم الحصول عليها دالة على المستوى الحقيقى للظاهرة المقيسة.

٣- الثبات يشير دائما إلى مستوى معين من الاتساق فى الدرجات، هذا المستوى من الاتساق مشروط بفترات زمنية، وعينات من الأداءات أو الأسئلة، ومفحوصين لهم خصائص محددة. فالثبات يختلف تبعا للفترات الزمنية بين مرات التطبيق المختلفة، ووفقا لخصائص الاختبار أو المقياس، ووفقا لخصائص عينة الأفراد.

٤- الثبات يعرَّف منطقيا على أنه نسبة التباين الحقيقى True variance إلى التباين العام total variance. والمقصود بالتباين هنا هو تباين الدرجات (القياسات) وليس تباين أداة القياس Measuring instrument، ويرجع السبب فى ذلك إلى أن القياس فى العلوم النفسية والتربوية والاجتماعية يتوقف على خصائص المجتمع الإحصائى الذى نقيسه باستخدام أداة القياس، وبناء على ذلك يظهر الثبات كمصطلح يشير إلى ثبات درجات الاختبار أو المقياس، وليس إلى ثبات الاختبار أو أداة القياس، فالاختبار الواحد يظهر قيما مختلفة من الثبات عندما يطبق على مجتمعات متباينة الخصائص مما يدل على أن الثبات يختص بالدرجة وليس بالأداة.

مما سبق يتضح أن ثبات أداة القياس يشير إلى الاتساق والدقة بالنسبة للنتائج التى يتم الحصول عليها من الأداة عندما يتم تكرار تطبيقها على نفس

الأفراد مع توحيد الشروط فى مرات التطبيق المختلفة، فالثبات يعنى دقة الاختبار فى القياس وعدم تناقضه مع نفسه. ويعتمد الثبات على مدى انحراف درجة كل فرد فى التطبيق الأول للاختبار عنها فى التطبيق الثانى لنفس الاختبار، كما يعتمد الثبات على الترتيب النسبى للأفراد فى المرات المختلفة لتطبيق الاختبار.

فاذا افترضنا أننا قمنا بتطبيق اختبار ما على عشرة طلاب، ثم قمنا بإعادة تطبيق نفس الاختبار مرة ثانية تحت نفس الشروط وبفارق زمنى أسبوع واحد بين التطبيقين، ثم قمنا بوضع النتائج فى الجدول (٢-١). وبالرجوع إلى مفهوم الثبات، فإنه يمكن أن نستخلص من الجدول الآتى:

جدول (٢-١) : درجات عشرة طلاب طبق عليهم أحد الاختبارات مرتين بفارق زمنى أسبوع بين التطبيقين

الطلاب	التطبيق الأول		التطبيق الثانى		المقارنة بين التطبيقين	
	الدرجة	الموقع النسبى	الدرجة	الموقع النسبى	انحراف الدرجة	الموقع النسبى
صلاح	٩٠	١	٩٥	١	٥	١
محمد	٨٥	٢	٨٩	٢	٤	٢
نصر	٨٤	٣	٨٥	٣,٥	١	٣، ٣,٥
عمرو	٨٣	٤	٨٥	٣,٥	٢	٤، ٣,٥
إبراهيم	٨٠	٥	٧٩	٥	١	٥
محمود	٧٥	٦,٥	٧٨	٦	٢	٦، ٦,٥
طه	٧٥	٦,٥	٧٤	٧	١	٦، ٧,٥
حسن	٧٤	٨	٧٣	٨	١	٨
رضوان	٧٣	٩	٧٠	٩	٣	٩
سالم	٧٠	١٠	٦٥	١٠	٥	١٠

١ - أن القيم المطلقة للدرجات قد تغيرت بين التطبيقين الأول والثانى، ولكنها تغيرات طفيفة (كما يظهر فى العمود السادس من الجدول).

٢ - أن الترتيب النسبى لمواقع الطلاب (المفحوصين) فى التطبيقين لم يتغير تغيرات ذات أهمية، وعليه فإنه يمكننا القول بأن هذا الاختبار يقيس الظاهرة بشكل دقيق ومتسق، وذلك لسببين هما:

الأول:

أن الدرجات التى حصلنا عليها من التطبيق الأول للاختبار تشابه تقريبا الدرجات التى حصلنا من التطبيق الثانى، حيث تشير هذه النتيجة إلى دقة الاختبار.

الثانى:

أن المواقع النسبية للطلاب العشرة لم تتغير فى المرتين تقريبا مما يعنى اتساق Consistancy النتائج التى يتم الحصول عليها نتيجة تطبيق الاختبار عبر فترتين زمنيتين مختلفتين.

أهمية الثبات:

يرتبط مصطلح «الثبات» بالعديد من المشكلات التى تكسبه أهميه خاصة تستلزم من الباحث أو القارئ المتخصص أن يلم ببعض الأفكار والمفاهيم الرياضية Mathematical ideas المرتبطة بهذا المصطلح لأنه عن طريق الإلمام بهذه المفاهيم يمكن وضع تعريف دقيق للثبات يجعلنا قادرين على إدراك أهميته وفهم الطرق المختلفة لاستخداماته. وعمومًا فإنه يمكن تحديد أهمية الثبات فى النقاط التالية:

١- الثبات أحد أهم المعايير (الشروط) الواجب توافرها فى أدوات القياس ولكونه يتأثر بعوامل الصدفة وأخطاء القياس، لذا يجب تقديرة قبل استخدام أدوات القياس للاستفادة منه كمؤشر لحجم الأخطاء التى يقع فيها الباحث وكذا تأثير عوامل الصدفة على عملية القياس، بالإضافة إلى موضوعية تقدير الدرجة على الاختبار أو المقياس.

٢ - لا يجوز أن يكـون الثبات بديلا عن الـصدق، فصدق الاخـتبار أو أداة القياس يأتى فى المرتبة الأولى، لكون الصدق يعد أول معايير (محكات) الجودة الواجب توافـرها فى أداة القياس حتـى يمكن إقرار صلاحيـتها، ولكن تظل أداة القياس تتمتع بالصـدق طالما كانت تستخدم فى حدود ما وضعـت من أجله، وبدون أى حـذف أو إضافة أو تـغييـر فى محـتوى بنودها، وأيضا دون تغـيير فى خصائص عينة الأفراد الـتى أعدت عليهم فى الأصل. إلا أن الأمـر يختلف بالـنسبة للثـبات فعلـى الرغم من أن الأداة التى تتمتع بالصدق تتمتع بثـبات النتائج، إلا أن تقدير الثبات يظل إجراء تستـلزمه التجارب البـحثية فى كل مـرة تطبق فيهـا أداة القياس، فالثبات بالنسبة لنفس الاختبار أو المقياس يتعرض لأخطاء تتعلق بالتطبيق وعوامل الصدفة (الـعوامل العشوائية)، هذه الأخطاء والعـوامل تختلف مـن باحـث لآخـر، ومن مـوقف لآخـر، ومن وقـت لآخـر، ومن مفحوص لآخر.

٣ - يتسم الثبات بالصبغة الإحصائية على عكس من الصدق، فقد يتم تقدير الصدق عن طريق التـحليل المنطقى، إلا أن التحليل المنطـقى لأى اختبار لا يعطى أى دليـل عملى عن الثبـات، ومن ثم أصبحت طـرق حساب الثبات تتـطلب فهم العديد من العمـليات الإحصائية التى أضـفت عليه أهمية خاصة.

٣ - مفهوم الثبات النسبى وطرق حسابه

يصنف الثبات وفقاً لطرق حسابه إلى نوعين رئيسيين هما:

- **الثبات النسبى** Relative reliability
- **الثبات المطلق** Absolute reliability

وفيما يلى شرح للثبات النسبى وطرق حسابه:

الثبات النسبى:

وهو ثبات يتأسس على مدى استقرار الأوضاع النسبية للأفراد داخل المجموعة أو العينة فى مرات التطبيق المختلفة، ويستخدم لحساب هذا النوع من الثبات أسلوبين من أساليب التحليل الإحصائى هما:

١- حساب الارتباط بين الفئات (مجموعات الدرجات) interclass Correlation ويطبق هذا الأسلوب بالنسبة لطرق: التجزئة النصفية، والصور المتكافئة وإعادة الاختبار.

٢- تحليل التباين ANOVA ويستخدم لحساب الارتباط داخل الفئات (مجموعات الدرجات) intraclass Correlation وهو أسلوب يستخدم فى الحالات التى يتكرر فيها الاختبار عدد من المحاولات، حيث تستخدم معادلتى كيودر - ريتشاردسون KR_{20} , kR_{21} ومعامل ألفا كرونباخ alpha.

ويستخدم لتقدير ثبات درجات الاختبار (المقياس) خمس طرق رئيسية هى:

١ - طريقة إعادة الاختبار Test-Retest Method.

٢ - طريقة الصور المتكافئة Equivalent - Forms Method.

٣ - طريقة التجزئة النصفية Split - Half Method.

٤ - طريقة الاتساق الداخلى Internal - Consistancy Method.

٥ - طريقة تحليل التباين Analysis of Variance Method.

ومن الملاحظ أن هذه الطرق تستخدم الدرجات الخام التى يتم الحصول عليها نتيجة تطبيق أدوات القياس المختلفة، وأن استخدام أيا من هذه الطرق يتوقف على ثلاثة شروط رئيسية هى:

أ - طبيعة الاختبار.

ب - إحصاء العينة وإحصاء المجتمع الأصلى.

جـ - الغرض الذى سوف يستخدم من أجله إحصاء العينة.

وفيما يلى شرح موجز لكل طريقة من هذه الطرق:

طريقة إعادة الاختبار Test - Retest Method

تستخدم طريقة إعادة الاختبار لتقويم الخطأ المرتبط بتطبيق الاختبار مرتين مختلفتين بالنسبة لوقت التطبيق، بمعنى أن الثبات فى هذه الطريقة يستخدم المعاينة الزمنية time sampling لتسجيل درجات الأداء على الاختبار أو المقياس فى فترتين زمنيتين.

وتعد هذه الطريقة من أبسط الطرق وأسهلها فى تقدير الثبات، فهى تقوم على أساس تطبيق الاختبار أو المقياس على مجموعة من الأفراد، ثم يعاد التطبيق مرة أخرى على نفس المجموعة، ثم يحسب معامل الارتباط بين التطبيقين لنحصل بذلك على معامل ثبات الاختبار أو المقياس، حيث يدل معامل الثبات المحسوب على معامل استقرار الاختبار stability coefficient.

ويمكن حساب معامل الارتباط بين التطبيقين بطريقة بيرسون أو سبيرمان ويتم حساب الدلالة الإحصائية له باستخدام الجداول الإحصائية الخاصة بمعاملات الارتباط، وإن كان يفضل استخدام الاغتراب كمؤشر لقوة معامل الارتباط الدال على معامل الاستقرار (الثبات).

ويوصى علماء القياس باستخدام طريقة إعادة الاختبار فى الحالات التالية:

١- فى حالة قياس السمات والخصائص التى يتوقع لها الاستقرار stable traits، بمعنى أنها لا تتغير تغيرات سريعة وحادة عبر الزمن، ومن أمثلة هذه السمات: الذكاء العام والقدرات البدنية والحركية.

٢- فى حالة الاختبارات غير المتجانسة heterogeneous وهى نمط من الاختبارات تستخدم لقياس متغيرات بدنية أو حركية أو نفسية مختلفة من

حيث كونها تتضمن أبعاداً أو أجزاءً مختلفة، ففى الاختبارات غير المتجانسة نجد أن أجزاء مختلفة من الاختبار تقيس أبعادًا (أجزاء) مختلفة من السمات أو الخصائص، لذلك لا يتوقع فى هذه الحالة الحصول على مؤشر مرتفع لثبات الاتساق الداخلى لدرجات الاختبار أو المقياس ، وبناء على ذلك يوصى علماء القياس باستخدام طريقة إعادة الاختبار Retest لحساب ثبات الاختبارات غير المتجانسة فى المحتوى، حيث تشير هذه الطريقة إلى ثبات الاختبار عن طريق إعادة تطبيقه the Retest reliability أو معامل إعادة الاختبار Retesting coefficient وهو ما يعرف باستقرار stability درجات الاختبار.

وتصنف اختبارات السرعة speed tests (اختبارات الورقة والقلم) كاختبارات غير متجانسة لكونها تتطلب الإجابة على مجموعة من الأسئلة خلال زمن محدد، حيث لا يستطيع كل المفحوصين الإجابة عن كل الأسئلة، وعليه يمكن النظر إلى الأسئلة التى تم الإجابة عليها على أنها غير متجانسة بالنسبة للأسئلة التى لم يتم الإجابة عليها تأسيساً على أن الزمن المخصص للاختبار يكون محدداً لدرجة صعوبته، بحيث لا يكون فى إمكان أيا من المفحوصين أن يجيب على كل الوحدات أو الفقرات أو الأسئلة فى الزمن المخصص للاختبار على الرغم من فرضية أن جميع الأسئلة متماثلة فى درجة صعوبتها.

ويوصى علماء القياس بعدم استخدام طريقة إعادة الاختبار فى الحالات التالية:

١- بالنسبة للاختبارات التى تقيس سمات أو صفات أو خصائص تكتسب بالتعلم وتتأثر بالمران والممارسة والتدريب مثل اختبارات المهارات skills الحركية فى الألعاب والأنشطة الرياضية.

٢- بالنسبة للاختبارات التى تقيس ظواهر تتأثر بالنضج الناشئ عن العمر الزمنى وبخاصة فى مراحل سن ٤ و ٥ سنوات مثل اختبارات الذكاء واختبارات النمو الجسمى والحالة الغذائية وبعض اختبارات الإدراك الحس - حركى والتوافقات.

٣- بالنسبة للاختبارات التى تقيس ظواهر تتأثر بالأحداث التى تقع بين التطبيقيين الأول والثانى كالقراءة ومشاهدة التليفزيون والتفاعل مع الأقران من أفراد الأسرة أو زملاء الدراسة أو العمل أو غيرها، مثل استبيانات استطلاعات الرأى بالنسبة لبعض القضايا والمشكلات الاقتصادية والاجتماعية والسياسية وغيرها.

٤- بالنسبة للاختبارات التى تقيس ظواهر تتميز بالتذبذب أو التأرجح Fluctuate بمرور الوقت، فقد أقر عدد من علماء النفس أن بعض مظاهر النزوع السلوكى behavioral disposions قد لا تظهر بشكل متسق طوال الوقت، فالشخص العدوانى قد لا يظهر نفس درجة العدوان طوال الوقت، وعلى هذا الأساس يقرر أتكنسن atkinson ١٩٨١م أنه لا يفضل حساب ثبات مقاييس العدوان والدافعية باستخدام طريقة إعادة الاختبار، وبالمثل فإنه لا يفضل على الإطلاق استخدام نفس الطريقة لحساب ثبات بعض القياسات الفسيولوجية كمعدل القلب مثلا.

٥- بالنسبة للاختبارات التى ينتقل فيها تأثير الأداء فى التطبيق الأول الى الأداء فى التطبيق الثانى carry - over effect ومن أمثلة ذلك اختبارات التحصيل والاختبارات المعرفية حيث يتوقع أن يتذكر المفحوصون إجاباتهم فى التطبيق الأول، وأن يعرفوا الإجابات الصحيحة لبعض الأسئلة خلال الفترة بين التطبيقين.

الفترة الزمنية بين التطبيقين الأول والثانى:

يشير جرونلند Gronlund ١٩٨١م إلى أن طريقة إعادة الاختبار تعنى إعطاء نفس الاختبار مرتين لنفس مجموعة الأفراد، على أن تكون الفترة الزمنية بين المرتين تتراوح بين عدد من الدقائق حتى عدد من السنوات، وعليه سوف تظل الفترة الزمنية المناسبة بين التطبيقين الأول والثانى، مثار جدل مستمر بين علماء القياس، فإذا كانت هذه الفترة قصيرة لمدة - يوم أو يومين - فإن اتساق النتائج سوف يزداد نتيجة تذكر التلاميذ لبعض إجاباتهم فى التطبيق الأول فينقلونها إلى التطبيق الثانى، هذا من ناحية، ومن ناحية أخرى فقد يحدث أن يؤثر التعب الناشئ عن التطبيق الأول - وبخاصة فى الاختبارات البدنية - تأثيراً سلبياً على الأداء فى التطبيق الثانى فينخفض اتساق النتائج (معامل الثبات).

وإذا كانت الفترة الزمنية بين التطبيقين طويلة - شهراً أو أكثر - فسوف يحدث الآتى:

- توقع حدوث تغيرات فى شروط وإجراءات التطبيق التى اتبعت فى المرة الأولى.
- توقع حدوث تغييرات فى سلوك المفحوصين خلال هذه الفترة الزمنية، وعليه فسوف يتوقع الحصول على معامل ثبات منخفض نتيجة طول الفترة الزمنية بين التطبيقين الأول والثانى.

وعموما فقد رأى علماء القياس أن الفترة الزمنية المناسبة بين التطبيقين تتوقف بشكل كبير على استخدامات النتائج التى سوف يتم الحصول عليها من الاختبار، فإذا كان الاختبار سوف يستخدم فى التنبؤ، حينئذ يشترط أن تتمتع درجات الاختبار بالاستقرار على مدى شهور قليلة حتى بضع سنوات، فاختبارات الاستعدادات البدنية تتمتع بثبات نسبى خلال نفس العام، وتصبح غير مستقرة إذا زادت الفترة عن العام.

وللخروج من هذه المشكلة الجدالية فقد أقر معظم المتخصصين فى القياس أن فترة أسبوعين بين التطبيقين الأول والثانى فترة مناسبة جدا لحساب الثبات بطريقة إعادة الاختبار لكونها ليست بالفترة القصيرة جداً أو الطويلة جداً.

طريقة الصور المتكافئة Equivalent - Forms Method:

يتطلب تقدير الثبات بطريقة الصور المتكافئة Parallel or equivalent forms استخدام صورتين متكافئتين من الاختبار، حيث يقصد بالتكافؤ هنا الآتى:

١- تساوى عدد الأسئلة فى الصورتين.

٢- تكافؤ درجة سهولة وصعوبة الأسئلة المتقابلة فى الصورتين (السؤال الأول فى الصورتين، والسؤال الثانى... إلخ).

٣- تساوى قيم معاملات الارتباط بين البنود فى كل من الصورتين (كل عبارة مع البعد الذى تنتمى إليه).

٤- تساوى قيم المتوسط الحسابى (س) والانحراف المعيارى (ع) لكلتا الصورتين ويتم استخدام هذه الطريقة **وفق الخطوات التالية:**

١- إعداد صورتين متكافئتين للاختبار إعدادا جيداً. وقد تكون لبعض الاختبارات صورتين (أو صيغتين) الصورة (أ) والصورة (ب).

٢- نقوم بتطبيق الصورتين المتكافئتين على مجموعة التقنين، وقد يتم تطبيق الصورتين معا فى نفس اليوم، وقد يتم تطبيق الصورة (أ) فى يوم والصورة (ب) بعد فترة زمنية مناسبة (يوم أو يومين أو ١٥ عشر يوما بحد أقصى).

٣- نقوم بحساب معامل الارتباط بين درجات مجموعة التقنين على الصورتين، حيث يدل معامل الارتباط المحسوب على معامل التكافؤ coefficient of equivalence والذى يعرف بعدة أسماء هى:

- معامل ثبات الصور المتبادلة Alternate - Forms Reliability
- معامل ثبات الصورة المتوازية Parallel - Forms Reliability
- معامل ثبات الصور المتساوية Comparable - forms Reliabitity

وعادة ما يستخدم معامل ارتباط بيرسون لحساب هذا المعامل، حيث يدلنا هذا المعامل على الآتى:

- يعد مقياساً للتكافؤ.
- يبين الدرجة التى تقيس بها كل من صورتى الاختبار نفس المظاهر أو الموضوعات المطلوب قياسها.
- إذا كان معامل الارتباط (الثبات) مرتفعاً دل ذلك على صلاحية الصورتين وحسن إعدادهما، وإذا جاء هذا المعامل منخفضاً دل ذلك على عدم صلاحية إحدى الصورتين أو كلاهما.

ومن الملاحظ أن طريقة الصور المتكافئة لا تدلنا على ثبات الخصائص المقيسة، وإنما تبين لنا إلى أى مدى يمثل الاختبار عينة من الخصائص والسمات ومظاهر السلوك المطلوب قياسها. فعند قياس التحصيل - على سبيل المثال - تظهر الآلاف من الأسئلة التى تستلزم أن يتضمنها الاختبار، ومع ذلك، ولأسباب تتعلق بالوقت والإمكانات المتاحة، يحدث أن يقتصر إعداد مثل هذه النوعية من الاختبارات على أعداد محدودة من الأسئلة التى يمكن استخدامها، هذه المجموعة من الأسئلة تمثل أفضل عينة من الأسئلة الممكنة لتمثيل المجال الذى تقيسه.

ويمكن النظر إلى طريقة الصور المتكافئة على أنها تطبيق للاختبار مرتين متتاليتين، حيث تختص كل مرة بإحدى صور الاختبار، ووفقا لهذا فإنه يمكن النظر إلى معامل الثبات المحسوب بطريقة الصورة المتكافئة على أنه مؤشر لاستقرار stability الاختبار وتكافؤ صورتيه equivalence فى ذات الوقت.

متى تستخدم طريقة الصور المتكافئة؟

تستخدم هذه الطريقة بشكل محدود فى مجال الاختبارات البدنية والحركية، لكونها خاصة باختبارات الورقة والقلم paper - and pencil tests التى تتطلب من المفحوص أن يعبر عن استجابته باستخدام الكتابة، وهى اختبارات تختلف تماماً عن الاختبارات العملية (الأدائية) performance tests التى تتطلب الاستجابة الحركية motor response وليس الاستجابة اللفظية (اللغوية) verbal response.

وعموما فإنه يمكن استخدام هذه الطريقة بالنسبة لاختبارات الورقة والقلم المقننة Standardized tests لكون معظم هذه الاختبارات له صورتان أو أكثر، من هذه الاختبارات على سبيل المثال: اختبارات التحصيل، والاختبارات المعرفية، وبعض القوائم والصور الخاصة بقياس الشخصية.

طريقة التجزئة النصفية Split - Half Method:

يمكن تقدير ثبات درجات الاختبار أو المقياس عن طريق التطبيق لمرة واحدة فقط، فى هذه الطريقة يطبق الاختبار أو المقياس على مجموعة واحدة من المفحوصين ثم يجزء الاختبار أو المقياس إلى نصفين متكافئين كالتالى:

١ - النصف الأول من العبارات فى مقابل النصف الثانى من العبارات:

الـ ٥٠ سؤالا الأولى فى مقابل الـ ٥٠ سؤالا الثانية

(فى حالة اختبار يتكون من ١٠٠ سؤال أو عبارة)

أو:

٢ - الأسئلة (العبارات) ذات الأرقام الفردية فى مقابل الأسئلة (العبارات) ذات الأرقام الزوجية.

٣ - وفق ما سبق يصبح لكل مفحوص مجموعتان من الدرجات الخام هما:

- مجموع درجات النصف الأول من العبارات (س)

- مجموع درجات النصف الثانى من العبارات (ص)

أو :

- مجموع درجات العبارات الفردية (س)

- مجموعة درجات العبارات الزوجية (ص)

٤ - يحسب معامل الارتباط بين (س، ص) لكل المفحوصين باستخدام معامل ارتباط بيرسون فنحصل على معامل ثبات نصف الاختبار وهو:

$\frac{١}{٢} . \frac{١}{٢}$ = معامل ثبات نصف الاختبار

ملحوظة مهمة:

تتطلب طريقة التجزئة النصفية تقسيم الاختبار إلى نصفين متساويين أو قريبين من التساوى، فإذا كان لدينا اختبار يتكون من (١٠) عبارات أو أسئلة، فإنه يمكن تقسيم هذا الاختبار إلى نصفين متساويين كالتالى:

أ- النصف الأول من العبارات = ٥ عبارات وهى العبارات أرقام (١، ٢، ٣، ٤، ٥) فى مقابل النصف الثانى من العبارات = ٥ عبارات وهى العبارات أرقام (٦، ٧، ٨ ، ٩ ، ١٠).

أو :

ب- العبارات الفردية = ٥ عبارات وهى العبارات أرقام (١، ٣، ٥، ٧، ٩) فى مقابل العبارات الزوجية = ٥ عبارات وهى العبارات أرقام (٢، ٤، ٦ ، ٨ ، ١٠).

obeikandl.com

أما إذا كان عدد العبارات ١١ عبارة، فإنه يمكن تقسيم هذا الاختبار إلى نصفين قريبين من التساوى nearly equal halves كالتالى:

أ- النصف الأول من العبارات = ٥ عبارات وهى العبارات أرقام (١، ٢، ٣، ٤، ٥)، فى مقابل النصف الثانى من العبارات = ٦ عبارات وهى العبارات أرقام (٦، ٧، ٨، ٩، ١٠، ١١).

أو:

ب- العبارات الفردية = ٦ عبارات وهى العبارات أرقام (١، ٣، ٥، ٧، ٩، ١١) فى مقابل العبارات الزوجية = ٥ عبارات وهى العبارات أرقام (٢، ٤، ٦، ٨، ١٠).

طرق تصحيح معامل ثبات نصف الاختبار:

يستخدم لتصحيح معامل ثبات نصف الاختبار للحصول على معامل ثبات الاختبار ككل أى من المعادلات التالية:

معادلة سبيرمان - براون: Spearman - Brown formula الصورة المختصرة وهى:

$$\text{معادلة (٢-١)} \qquad \overset{\frown}{\text{١٠١}} = \frac{\overset{\frown}{\frac{١}{٢} \cdot \frac{١}{٢}} \times ٢}{١ + \overset{\frown}{\frac{١}{٢} \cdot \frac{١}{٢}}}$$

حيث إن:

ρ_{11} = معامل ثبات كل الاختبار.

$\rho_{\frac{1}{2}\frac{1}{2}}$ = معامل الارتباط بين نصفى الاختبار (معامل ثبات نصف الاختبار).

٢ = مقدار ثابت.

١ = مقدار ثابت.

مثال: (٢ - ١)

احسب معامل ثبات الاختبار ككل إذا كان معامل الارتباط بين نصفى الاختبار يساوى ٠,٧٠

بتطبيق المعادلة (٢ - ١) ينتج أن:

$$\rho_{11} = \frac{٢ \times ٠,٧٠}{١ + ٠,٧٠}$$

$$= \frac{١,٤}{١,٧٠}$$

$$= ٠,٨٢$$

وتستخدم هذه المعادلة فى اختبارات التحصيل والقدرات، حيث يشترط بالنسبة لاستخدامها أن تكون تباينات variances نصفى الاختبار متساوية.

معادلة رولون The Rulon Formula **وهى:**

$$\text{ر}_{\text{ن ن}} = 1 - \frac{\text{ع}_{\text{د}}^{2}}{\text{ع}_{\text{ت}}^{2}}$$

معادلة (٢ - ٢)

حيث إن:

$\text{ر}_{\text{ن ن}}$ = معامل ثبات كل الاختبار.

$\text{ع}_{\text{د}}^{2}$ = تباين الفرق بين درجات المفحوصيين فى النصف الأول ودرجاتهم فى النصف الثانى من الاختبار.

$\text{ع}_{\text{ت}}^{2}$ = تباين الاختبار ككل.

مثال: (٢ - ٢)

طبق اختبار يتكون من ١٢ عبارة على ١٠ مفحوصيين فكان مجموع درجات العبارات الفردية، ومجموع العبارات الزوجية لكل مفحوص كالتالى:

المفحوص:	أ	ب	جـ	د	هـ	و	ز	ح	ط	ى
مجموع العبارات الفردية:	٦	٦	٢	٣	٤	٢	٤	٦	٥	٥
مجموع العبارات الزوجية:	٥	٥	٢	٢	٣	٠	٤	٦	٦	٤

obeikandi.com

والمطلوب:

حساب معامل ثبات كل الاختبار باستخدام معادلة «رولون».

الحـــل: نقـــوم بعمل جدول إحـصائى متـضمنًا العـمليات الإحصـائية الأساسية كالتالى:

المفحوص	قيم العبارات الفردية $س_د$	قيم العبارات الزوجية $س_ز$	الفرق $ع_د$ $(س_د - س_ز)$	مربع الفرق $ع_د^2$	قيم جميع العبارات س	مربع $س^2$	مربع $س_د^2$	مربع $س_ز^2$
أ	٦	٥	+١	١	١١	١٢١	٣٦	٢٥
ب	٦	٥	+١	١	١١	١٢١	٣٦	٢٥
جـ	٢	٢	٠	٠	٤	١٦	٤	٤
د	٣	٢	+١	١	٥	٢٥	٩	٤
هـ	٤	٣	+١	١	٧	٤٩	١٦	٩
و	٢	٠	+٢	٤	٢	٤	٤	٠
ز	٤	٤	٠	٠	٨	٦٤	١٦	١٦
ح	٦	٦	٠	٠	١٢	١٤٤	٣٦	٣٦
ط	٥	٦	-١	١	١١	١٢١	٢٥	٣٦
ى	٥	٤	+١	١	٩	٨١	٢٥	١٦
	مجـ $س_د$ =	مجـ $س_ز$ =	مجـ $ع_د$ =	مجـ $ع_د^2$ =	مجـ س =	مجـ $س^2$ =	مجـ $س_د^2$ =	مجـ $س_ز^2$ =
	٤٣	٣٧	٦	١٠	٨٠	٧٤٦	٢٠٧	١٧١

يلاحظ من العمليات الإحصائية التى يتضمنها الجدول السابق أن:

مجموع قيم العبارات الفردية مجـ $س_د$ = ٤٣ درجة

ومجموع قيم العبارات الزوجية مجـ $س_ز$ = ٣٧ درجة

ومجموع الفروق مجـ $ع_د$ = ٦

ومجموع مربع الفروق مجـ $ع_د^2$ = ١٠

ومجموع قيم جميع العبارات مجـ س = ٨٠

ومجموع مربعات جميع القيم مجـ $س^2$ = ٧٤٦

ومجموع مربعات المحاولات الفردية مجـ $س_د^2$ = ٢٠٧

ومجموع مربعات المحاولات الزوجية مجـ $س_ز^2$ = ١٧١

ولحساب تباين الفرق $ع_د^2$ تطبق المعادلة التالية:

$$ع_د = \sqrt{\frac{١}{ن} \text{ن مجـ } ع_د^2 - (\text{مجـ } ع_د)^2}$$ معادلة (٢ - ٣)

وبالتعويض فى هذه المعادلة ينتج أن:

$$ع_د = \sqrt{\frac{١}{١٠} [١٠ \times ١٠ - (٦)^2]}$$

$$= \sqrt{\frac{١}{١٠} (١٠٠ - ٣٦)}$$

$$= ٠,٨$$

إذا $ع_د^2$ $= (٠,٨)^2 = ٠,٦٤$

obeikandl.com

(حيث إن التباين = مربع الانحراف المعيارى)

وبالمثل يتم حساب التباين الكلى للاختبار $ع_{ت}^{٢}$ بتطبيق المعادلة التالية:

$$ع_{ت} = \frac{١}{ن} \sqrt{ن \text{ مجـ س}^{٢} - (\text{مجـ س})^{٢}}$$ معادلة (٢ - ٤)

وبالتعويض فى هذه المعادلة ينتج أن:

$$ع_{ت} = \frac{١}{١٠} \sqrt{[١٠ \times ٧٤٦ - (٨٠)^{٢}]}$$

$$= \frac{١}{١٠} \times ٣٢,٥٥٧٦$$

$$= ٣,٢٦$$

$$\text{إذا } ع_{ت}^{٢} = (٣,٢٦)^{٢} = ١٠,٦٣$$

(حيث إن التباين = مربع الانحراف المعيارى)

أسفرت العمليات الإحصائية السابقة عن الآنى:

$$ع_{ت}^{٢} = ٠,٦٤$$

$$ع_{ت}^{٢} = ١٠,٦٣$$

حينئذ يمكن حساب معامل ثبات كل الاختبار بتطبيق معادلة رولون كالتالى:

$$\text{ر}_{١٠١} = ١ - \frac{٠,٦٤}{١٠,٦٣}$$

$$= ١ - ٠,٠٦$$

$$= ٠,٩٤$$

معادلة فلانجان Falngan Formula **وهى:**

$$\text{ر}_{\text{١٠١}} = ٢ \left(١ - \frac{(\text{ع}_{\text{١}}^{٢} - \text{ع}_{\text{٢}}^{٢})}{\text{ع}_{\text{ت}}^{٢}}\right) \qquad \text{معادلة (٢ - ٥)}$$

حيث إن:

$\text{ر}_{\text{١٠١}}$ = معامل ثبات كل الاختبار.

$\text{ع}_{\text{١}}^{٢}$ = تباين درجات النصف الأول.

$\text{ع}_{\text{٢}}^{٢}$ = تباين درجات النصف الثانى للاختبار.

$\text{ع}_{\text{ت}}^{٢}$ = التباين الكلى للاختبار.

ويمكن تطبيق معادلة فلانجان على البيانات المعطاة فى المثال السابق رقم (٢ - ٢) لحساب ثبات كل الاختبار وفق الخطوات التالية:

أولًا: حساب تباين درجات النصف الأول للاختبار ($\text{ع}_{\text{١}}^{٢}$) بتطبيق المعادلة (٢ - ٤) وهى:

$$\text{ع}_{\text{١}} = \sqrt{\text{ن مجـ س}^{٢} - (\text{مجـ س})^{٢}}$$

حيث أن:

ن = ١٠

مجـ $\text{س}^{٢}$ = ٢٠٧ (للعبارات الفردية)

مجـ س = ٤٣ (مجموع قيم العبارات الفردية)

وبالتعويض فى المعادلة رقم (٢-٤) ينتج أن:

obeikandl.com

$$ع_{١} = \frac{١}{١٠} \sqrt{١٠ \times ٢٠٧ - (٤٣)^{٢}}$$

$$= \frac{١}{١٠} \sqrt{٢٠٧٠ - ١٨٤٩}$$

$$= ١,٤٩$$

إذاً $ع_{١}^{٢}$ $= (١,٤٩)^{٢} = ٢,٢٢$

ثانيا : حساب تباين درجات النصف الثانى للاختبار ($ع_{٢}^{٢}$) بتطبيق نفس المعادلة (٢ - ٤) حيث إن :

ن = ١٠

مجـ س٢ = ١٧١ (للمحاولات الزوجية)

مجـ س = ٣٧ (مجموع قيم العبارات الزوجية)

وبالتعويض فى المعادلة (٢ - ٤) ينتج أن:

$$ع_{٢}^{٢} = \frac{١}{١٠} \sqrt{١٠ \times ١٧١ - (٣٧)^{٢}}$$

$$= \frac{١}{١٠} \sqrt{١٧١٠ - ١٣٦٩}$$

$$= ١,٨٥$$

إذا $ع_{٢}^{٢}$ $= (١,٨٥)^{٢} = ٣,٤٢$

ثالثا : لحساب التباين الكلى للاختبار ($ع_{ت}^{٢}$) بتطبيق نفس المعادلة (٢ - ٤) حيث يظهر أن :

$ع_{ت}^{٢}$ = ١٠,٦٣ (من تطبيق معادلة رولون)

رابعا : يتم حساب ثبات كل الاختبار بتطبيق معادلة فلانجان (٢-٥) كالتالى :

$$\text{ر}_{١٠١} = ٢ (١ - \frac{٢,٢٢ + ٣,٤٢}{١٠,٦٣})$$

$$= ٢ (١ - \frac{٥,٦٤}{١٠,٦٣})$$

$$= ٠,٩٣٩$$

إذاً $\text{ر}_{١٠١} = ٠,٩٤$

(وهى نفس النتيجة التى تم الحصول عليها من معادلة رولون)

مما سبق يلاحظ الآتى :

أ - يتطلب استخدام معادلة رولون توافر نوعين من البيانات هما :

١ - تباين الفرق بين درجات المفحوصين على نصفى الاختبار.

٢ - التباين الكلى للاختبار.

ب- ويتطلب استخدام معادلة فلانجان ثلاثة أنواع من البيانات هى :

١ - تباين درجات النصف الأول للاختبار.

٢ - تباين درجات النصف الثانى للاختبار.

٣ - تباين درجات كل الاختبار.

ومن الملاحظات الجديرة بالاهتمام أنه يمكن استخدام طريقة التجزئة النصفية لحساب ثبات درجات بعض الاختبارات الحركية والقياسات الجسمية التى يمكن تطبيقها لعدد من المحاولات أو المرات (القياسات) المتتالية دون أن يتأثر الأداء بالتعب أو بانتقال أثر التدريب.

obeikandl.com

ومن أمثلة هذه الاختبارات:

١ - **اختبارات القدرة العضلية** Muscular power (القوة السريعة أو المتفجرة) مثل اختبارات: الوثب العمودى vertical Jump ، والوثب العريض من الثبات standing broad jump، ودفع الكرة الطبية ٣كجم Medicine Ball Put وغيرها، حيث تسمح شروط الاختبار بإعطاء المختبر (المفحوص) ثلاث أو أربع محاولات متتالية فى نفس الوقت.

٢ - **اختبارات التوافق والدقة** وبعض الاستجابات الحركية الانتقائية كما فى حالة اختبار نيلسون للاستجابة الحركية الانتقائية Nelson choice Response movement حيث يعطى المختبر ١٠ محاولات متتالية بين كل محاولة والأخرى ٢٠ ثانية، واختبار نيلسون للسرعة الحركية Nelson speed of movement حيث يعطى المختبر ٢٠ محاولة متتالية، وغيرها من الاختبارات[١].

٣ - **اختبارات المهارات فى الألعاب الرياضية**، حيث تسمح الشروط الخاصة بمعظم هذه الاختبارات بتنفيذ الاختبار لعدد من المحاولات المتتالية مرة واحدة[٢].

٤ - **القياسات الجسمية (المورفولوجية)** كما فى حالة قياس أطوال أو محيطات أو اتساعات بعض أجزاء الجسم حيث يفضل إجراء القياس الواحد لعدد ٣ أو ٤ مرات متتالية[٣].

---

(١) انظر هذه الاختبارات وغيرها فى المرجع الآتى:
محمد حسن علاوى ومحمد نصرالدين رضوان: اختبارات الأداء الحركى، دار الفكر العربى، ٢٠٠١م.

(٢) انظر هذه الاختبارات وغيرها فى المراجع التالية :
محمد حسن علاوى ومحمد نصر الدين رضوان: الاختبارات المهارية والنفسية فى المجال الرياضى، دار الفكر العربى، ١٩٨٧م.

(٣) محمد نصر الدين رضوان: المرجع فى القياس الجسمية، دار الفكر العربى، ١٩٩٧م.

مثال : (٢ - ٣)

طبق اختبار دفع الكرة الطبية (٣ كجم) على خمسة لاعبين، وكان كل لاعب يقوم بتنفيذ الاختبار عدد ٦ محاولات متتالية، فكانت النتائج مسجلة بالمتر كالتالى:

المحـــــاولات

اللاعب	١	٢	٣	٤	٥	٦
أ	١٠	١١	١١	١٢	١١	١٢
ب	١٤	١٣	١٤	١٤	١٥	١٤
جـ	١٢	١٢	١٢	١٣	١٢	١٣
د	٩	٨	٨	٩	٩	٨
هـ	٧	٩	٨	٨	٩	٩

والمطلوب :

أولا : حساب معامل ثبات الاختبار ككل باستخدام معادلة سبيرمان - براون.

ثانيا : حساب معامل ثبات الاختبار ككل باستخدام معادلة كل من: رولون، وفلانجان.

الحـــل :

أولا : يتم حساب ثبات الاختبار ككل باستخدام معادلة سبيرمان - براون كالتالى:

١ - تقسم محاولات الاختبار إلى نصفين، حيث يمثل النصف الأول مجموع نتائج المحاولات الفردية، ويمثل النصف الثانى مجموع نتائج المحاولات الزوجية لكل لاعب على حدة كالتالى:

اللاعب	مجموع المحاولات الفردية	مجموع المحاولات الزوجية
أ	٣٢	٣٥
ب	٤٣	٤١
جـ	٣٦	٣٨
د	٢٦	٢٥
هـ	٢٤	٢٦

٢ - مما سبق يتبين أن درجات اللاعبين فى المحاولات الست للاختبار أصبحت تتكون من مجموعتين، المجموعة الأولى وتضم درجات المحاولات الفردية (س)، والمجموعة الثانية وتضم درجات المحاولات الزوجية (ص)، وبتطبيق معامل ارتباط بيرسون للدرجات الخام، فإنه يمكن حساب معامل الارتباط بين نصفى الاختبار، حيث نجد أن هذا المعامل يساوى ٠,٨٩ أى أن:

$$\text{(معامل ثبات نصف الاختبار)} = ٠,٨٩ \quad \overbrace{\frac{١}{٢} \cdot \frac{١}{٢}}$$

وبتطبيق معادلة سبيرمان - براون (معادلة ٢ - ١) لتصحيح معامل ثبات نصف الاختبار والحصول على معامل ثبات الاختبار ككل ينتج أن:

$$\overbrace{\quad}^{١٠١} = \frac{٢ \times ٠,٨٩}{١ + ٠,٨٩}$$

$$= \frac{١,٨٧}{١,٨٩}$$

$$= ٠,٩٤$$

إذا معامل الثبات لجميع المحاولات (٦ محاولات) هو ٠,٩٤.

ثانيا: لحساب معامل ثبات الاختبار ككل باستخدام معادلة رولون، ومعادلة فلانجان، تستخدم الخطوات التالية:

١ - نقوم بـوضع نتائج المحـاولات الفردية والمحـاولات الزوجيـة فى جدول إحصائى كالتالى:

المفحوص	قيم المحاولات الفردية $س_د$	قيم المحاولات الزوجية $س_ز$	الفرق $ع_د$ $(س_د - س_ز)$	مربع الفرق $ع_د$	قيم جميع المحاولات س	مربع القيم $س^2$	مربع $س^2_د$	مربع $س^2_ز$
أ	٣٢	٣٥	-٣	٩	٦٧	٤٤٨٩	١٠٢٤	١٢٢٥
ب	٤٣	٤١	+٣	٤	٨٤	٧٠٥٦	١٨٤٩	١٦٨١
جـ	٣٦	٣٨	-٢	٤	٧٤	٥٤٧٦	١٢٩٦	١٤٤٤
د	٢٦	٢٥	+١	١	٥١	٢٦٠١	٦٧٦	٦٢٥
هـ	٢٤	٢٦	-٢	٤	٥٠	٢٥٠٠	٥٧٦	٦٧٦
	مجـ $س_د$ = ١٦١	مجـ $س_ز$ = ١٦٥	مجـ $ع_د$ = -٤	مجـ $ع^2_د$ = ٢٢	مجـ س = ٣٢٦	مجـ $س^2$ = ٢٢١٢٢	مجـ $س^2_د$ = ٥٤٢١	مجـ $س^2_ز$ = ٥٦٥١

٢ - نقوم بإجراء العمليات الإحصائية وفق ما هو مبين بالجدول السابق حيث يلاحظ أن:

مجـ $س_د$ = ١٦١

مجـ $س_ز$ = ١٦٥

مجـ $ع_د$ = -٤

مجـ $ع^2_د$ = ٢٢

مجـ س = ٣٢٦

مجـ $س^2$ = ٢٢١٢٢

مجـ $س^2_د$ = ٥٤٢١

مجـ $س^2_ز$ = ٥٦٥١

obeikandl.com

٣ - يتم حساب تباين الفرق ($ع_د^2$) بتطبيق المعادلة (٢ - ٣) كالتالى :

$$ع_د = \frac{1}{5} \sqrt{5 \times 22 - (-4)^2}$$

$$= ٠,٢ \sqrt{١١٠ - ١٦}$$

$$= ٠,٢ \times ٩٤$$

$$= ٠,٢ \times ٩,٧٠$$

$$= ١,٩٣٩$$

إذًا $ع_د^2 = (١,٩٣٩)^2 = ٣,٧٦$

٤ - وبالمثل يتم حساب التباين الكلى للاختبار $ع_ت^2$ بتطبيق المعادلة (٢-٤) كالتالى:

$$ع_ت = \frac{1}{5} \sqrt{5 \times 22122 - (326)^2}$$

$$= ٠,٢ \sqrt{١١٠٦١٠ - ١٠٦٢٧٦}$$

$$= ٠,٢ \sqrt{٤٣٣٤}$$

$$= ٠,٢ \times ٦٥,٨٣٣$$

$$= ١٣,١٦٧$$

إذًا $ع_ت^2 = (١٣,١٦٧)^2 = ١٧٣,٣٦$

obeikandl.com

٥ – بالتعويض فى معادلة رولون ينتج أن

$$\text{ر}_{١٠١} = ١ - \frac{٣,٧٦}{١٧٣,٣٦}$$

$$= ١ - ٠,٠٢$$

$$= ٠,٩٨$$

وبالمثل يمكن تطبيق معادلة فلانجان وفق الخطوات التالية:

١ – حساب تباين درجات النصف الأول من للاختبار ($ع^{٢}_{١}$) بتطبيق المعادلة (٢ – ٤) كالتالى :

$$ع_{١} = \frac{١}{٥} \sqrt{٥ \times ٥٤٢١ - (١٦١)^{٢}}$$

$$= ٠,٢ \sqrt{٢٧١٠٥ - ٢٥٩٢١}$$

$$= ٠,٢ \sqrt{١١٨٤}$$

$$= ٠,٢ \times ٣٤,٤١ = ٦,٨٨$$

$$\text{إذًا } ع^{٢}_{١} = (٦,٨٨)^{٢} = ٤٧,٣٣$$

٢– حساب تباين درجات النصف الثانى للاختبار ($ع^{٢}_{٢}$) بتطبيق نفس المعادلة (٢–٤) كالتالى :

$$= \frac{١}{٥} \sqrt{٥ \times ٥٥٦٥١ - (١٦٥)^{٢}}$$

$$= ٠,٢ \sqrt{٢٨٥٥ - ٢٧٢٥}$$

$$= ٠,٢ \sqrt{١٠٣٠}$$

obeikandl.com

$= ٠,٢ \times ٣٢,٠٩$

$= ٦,٤٢$

إذا $ع_{١}^{٢} = (٦,٤٢)^{٢} = ٤١,٢٢$

٣ - حساب التباين الكلى للاختبار ($ع_{ن}^{٢}$) وقد بلغ ١٧٣,٣٦ (الخطوة رقم: ٣ من إجراءات تطبيق معادلة رولون).

٤ - وبتطبيق معادلة فلانجان معادلة (٢-٥) على هذه المجموعة من المعطيات ينتج أن :

$$ر_{١٠١} = ٢ (١ - \frac{٤١,٢٢ + ٤٧,٣٣}{١٧٣,٣٦})$$

$$= ٢ (١ - \frac{٨٨,٥٥}{١٧٣,٣٦})$$

$$= ٢ (١ - ٠,٥١)$$

$$= ٢ \times ٠,٤٨٩$$

$= ٠,٩٨$ (وهى نفس النتيجة التى تم الحصول عليها من معادلة رولون)

مميزات طريقة التجزئة النصفية:

تمتاز طريقة التجزئة النصفية بالآتى:

١- تستخدم عندما يتعذر استخدام أيا من الطريقتين السابقتين (إعادة الاختبار أو الصور المتكافئة).

٢- تعد مناسبة لبعض الاختبارات التى تتأثر بشكل حساس بالتعليم وانتقال أثر التدريب أو التعب الذى يحدث نتيجة التطبيق الأول ويمكن أن ينتقل تأثيره إلى التطبيق الثانى.

obeikan.com

٣ - تطبيق الاختبار مرة واحدة بدلاً من مرتين أو إعداد صورة متكافئة للاختبار، يعمل على توفير المجهود والوقت والتغلب على ظاهرة تسرب بعض المفحوصين (المختبرين) فى التطبيق الثانى.

٤ - تطبيق الاختبار لمرة واحدة فقط يجنبنا جدلية الفترة الزمنية الفاصلة بين التطبيقين الأول والثانى والتى ما تزال محل اهتمام علماء القياس.

٥ - تمكننا من التغلب على المشكلات المتعلقة باختلاف شروط التطبيق والتى يحتمل أن تحدث فى المرة الثانية للتطبيق، وأيضا التغلب على التباين المحتمل حدوثه بالنسبة للتعليمات المعطاة للمختبرين عندما يطبق عليهم الاختبار مرتين.

٦- تصنف طريقة التجزئة النصفية كأحد مقاييس الاتساق الداخلى Measure of internal consistency لكونها تتعامل مع جميع مفردات الدرجات مقسمة على قسمين أو نصفين.

المآخذ على هذه الطريقة:

١ - قد يتعرض المفحوصين لعوامل الإجهاد والملل والضيق عندما تطبق اختبارات أو قوائم تحتوى على عدد كبير من الأسئلة مثلا (١٢٠ سؤال أو عبارة)، فإذا أخذنا البنود أرقام من (١ - ٦٠) ممثلة للنصف الأول من الاختبار، والبنود من (٦١ - ١٢٠) ممثلة للنصف الثانى من الاختبار، فى هذه الحالة نجد أن إجابات المفحوصين على النصف الثانى من الاختبار سوف لا تعطى نتائج يوثق بها، وذلك لسببين هما:

أ- أن تركيز الاهتمام فى الإجابة على عبارات النصف الأول قد يستهلك جهد وطاقة المفحوص.

ب- أن ضيق الوقت أو التعب والملل أو غيرهما قد تجعل الاهتمام بالإجابة على عبارات النصف الثانى أقل منه بالنسبة للنصف الأول من الاختبار.

٢- أنه فى حالة الأخذ بأسلوب التجزئة النصفية للاختبار على أساس العبارات الفردية odd فى نصف، والعبارات الزوجية even فى النصف الآخر، فى هذه الحالة فإنه يحتمل أن نحصل على تباينات غير متساوية لنصفى الاختبار لكون العبارات فى النصفين ليست على درجة واحدة من الصعوبة.

٣- قد تتضمن بعض الاختبارات أو القوائم الأسئلة (البنود أو العبارات) بدون ترتيب محدد كقوائم الشخصية واختبارات السرعة مثلا، مما يمثل تحديا لهذه الطريقة، لذا ينصح بعض المتخصصين القيام بترتيب الأسئلة فى مثل هذه الحالة حسب درجة صعوبتها ثم إعطائها أرقاماً مسلسلة بعد ذلك.

٤- يقتصر بعض الباحثين عند حساب الثبات بطريقة التجزئة النصفية على تقدير معامل الارتباط بين نصفى الاختبار فقط، دون تطبيق معادلة سبيرمان - براون لتصحيح هذا المعامل، وهذا الخطأ الشائع يجب التنبه له لكون معامل الارتباط بين نصفى الاختبار يشير فقط إلى ثبات نصف الاختبار.

والجدول (٢-٢) يبين تأثير استخدام معادلة سبيرمان - براون على معاملات الارتباط بين نصفى الاختبار.

جدول (٢-٢)

يوضح تأثير استخدام معادلة سبيرمان - براون لتصحيح معاملات الارتباط بين نصفى الاختبار

معاملات الارتباط بين نصفى الاختبار	معاملات الارتباط بعد التصحيح	مقدار التغير
٠,٠٥	٠,٠٩	٠,٠٤
٠,١٥	٠,٢٦	٠,١١
٠,٢٥	٠,٤٠	٠,١٥
٠,٣٥	٠,٥٢	٠,١٧
٠,٤٥	٠,٦٢	٠,١٧
٠,٥٥	٠,٧١	٠,١٦
٠,٦٥	٠,٧٩	٠,١٤
٠,٧٥	٠,٨٦	٠,١١
٠,٨٥	٠,٩٢	٠,٠٧
٠,٩٥	٠,٩٧	٠,٠٢

obeikandl.com

طريقة الاتساق الداخلى Internal Consistency Method

توجد بالإضافة إلى طريقة التجزئة النصفية مجموعة أخرى من الطرق لتقدير الاتساق الداخلى للاختبار. ففى عام ١٩٣٧م قدم كيودر وريتشاردسون Kuder and Richardson إنجازهما المتميز لتقويم الثبات من خلال تطبيق الاختبار مرة واحدة دون اللجوء إلى تقسيم الاختبار إلى نصفين بأية طريقة تحكمية، وأيضا للتغلب على مشكلة عدم تساوى التباينات بين نصفى الاختبار والتى تعد من أهم متطلبات طريقة التجزئة النصفية.

وتُعنى طريقة الاتساق الداخلى بالآتى:

١- مدى ارتباط البنود أو الفقرات أو الوحدات مع بعضها البعض داخل الاختبار أو المقياس.

٢- مدى ارتباط كل بند أو فقرة أو وحدة مع الاختبار ككل.

٣- التحقق من مدى الاتساق فيما بين البنود أو الفقرات أو الوحدات لكون هذا الاتساق يتأثر بمصدرين للخطأ هما:

أ) أخطاء محتوى البنود (الفقرات أو الوحدات).

ب) أخطاء عدم تجانس البنود (الفقرات أو الوحدات).

ويشكل التجانس homogeneus موضوع الاهتمام الرئيسى بالنسبة للاتساق الداخلى، حيث تستهدف طريقة الاتساق الداخلى البحث عن الدرجة التى يمكن الوثوق فيها كمؤشر لثبات الاتساق الداخلى للاختبار أو المقياس Internal-consistency reliability، حيث يعرف هذا بالثبات الداخلى لأداة القياس.

معادلات كيودر - ريتشاردسون :

وضع كيودر وريتشاردسون معادلتين ٢٠ و ٢١ واللتين حازتا على شهرة فائقة وأصبحتا تعرف بـ K-R20 وهى المعادلة الرئيسية، والمعادلة الثانية

وتعرف بـ K-R21 وهى مشتقة - مع التعديل - من المعادلة الرئيسية K-R20.

المعادلة K-R20:

وصورتها الرياضية كالتالى:

$$\text{ر}_{\text{ت ت}} = \frac{\text{ن}}{\text{ن} - ١} \left(\frac{\text{ع}^{٢} - \text{مجـ ب ك}}{\text{ع}^{٢}} \right)$$ معادلة (٢ ـ ٦)

حيث إن:

$\text{ر}_{\text{ت ت}}$ = معامل الثبات المقدر من المعادلة K-R20

$\text{ع}^{٢}$ = التبايـن الكلى لدرجات الاختبار

ب = نسبة الذيـن أجابوا إجابات صحيحـة على كل وحدة من وحدات الاختبار.

ك = نسبة الذين أجابوا إجابات خاطئة على كل وحدة من وحدات الاختبار حيث ك بالنسبة لكل وحدة تساوى (١- ب).

ن = عدد بنود أو وحدات الاختبار.

مجـ ب ك = مجموع حاصل ضرب نسبة الإجابات الصحيحة × نسبة الإجابات الخاطئة.

ويشترط لاستخدام هذه المعادلة الآتى:

١ - أن تكون كل بنود أو وحدات أو فقرات الاختبار متجانسة، بمعنى أن تقيس كل فقرة أو عبارة نفس البعد أو المحور الذى تنتمى إليه.

٢- أن تكون معاملات الارتباط البينية للفقرات المندرجة تحت البعد الواحد غير صفرية.

٣- أن تكون الدرجة المعطاة لكل فقرة أو عبارة إما (١ أو صفر) حيث تمنح الدرجة (١) للإجابة الصحيحة، والدرجة (صفر) للإجابة الخطأ. وهذا

obeikandl.com

يعنى أن معادلة كيودر - ريتشاردسون تستخدم فى حالة الاختبارات ثنائية البعد، والتى تدور الاستجابة فيها حول احتمالين اثنين فقط إما صح (✓) وإما خطأ (×).

٤- ألا تتباين معاملات صعوبة الفقرات تباينا كبيراً، فى حين يشترط فى تطبيق المعادلة K-R21 أن يكون لكل الفقرات نفس درجة الصعوبة، لأنه فى حاله تباين درجات صعوبة الفقرات تباينًا كبيراً فإن الثبات المحسوب بالمعادلة K-R21 يكون منخفضا.

مثال : (٢-٤)

طبق اختبار يتكون من ٦ فقرات على ٥٠ تلميذاً، وكانت الإجابة على الفقرات وفقاً لمقياس تقدير ثنائى (صح أو خطأ). وقد جاءت الإجابات على الاختبار كالتالى:

الوحدات	: الأولى	الثانية	الثالثة	الرابعة	الخامسة	السادسة
الإجابة الصحيحة	: ١٢	٤١	١٨	٢٩	٣٠	٤٧

والمطلوب :

حساب معامل ثبات هذا الاختبار باستخدام معادل K-R20 .

الحل :

١ - نقوم بإيجاد قيمة (ب) وذلك بقسمة عدد الإجابات الصحيحة لكل فقرة على العدد الكلى للمفحوصين الذين طبق عليهم الاختبار كالتالى:

$\frac{١٢}{٥٠}$ = ٠,٢٤ ، $\frac{٤١}{٥٠}$ = ٠,٨٢ ، $\frac{١٨}{٥٠}$ = ٠,٣٦ وهكذا

حتى الفقرة السادسة

٢ - نقوم بإيجاد قيمة (ك) لكل فقرة بطرح قيمة ب (فى الخطوة السابقة) من ١ (واحد)، فنحصل بذلك نسب الإجابات غير الصحيحة لكل فقرة كالتالى:

١ - ٠,٢٤ = ٠,٧٦

١ - ٠,٨٢ = ٠,١٨

١ - ٠,٣٦ = ٠,٦٤

. . . . . .

وهكذا حتى الفقرة السادسة

٣ - نقوم بإيجاد (ب × ك) لكل فقرة بضرب الناتج فى الخطوتين (١,٢) كالتالى:

٠,٢٤ × ٠,٧٦ = ٠,١٨

٠,٨٢ × ٠,١٨ = ٠,١٥

٠,٣٦ × ٠,٦٤ = ٠,٢٣

. . . . . .

وهكذا حتى الفقرة السادسة

٤ - نقوم بحساب مجـ ب ك بجمع القيم المحسوبة من الخطوة ٣.

٥ - توضع النتائج السابقة فى جدول إحصائى كالتالى:

الفقرات	الإجابات الصحيحة (س)	(ب) (من الخطوة١)	ك (من الخطوة٢)	ب × ك (من الخطوة٣)
١	١٢	٠,٢٤	٠,٧٦	٠,١٨
٢	٤١	٠,٨٢	٠,١٨	٠,١٥
٣	١٨	٠,٣٦	٠,٦٤	٠,٢٣
٤	٢٩	٠,٥٨	٠,٤٢	٠,٢٤
٥	٣٠	٠,٦٠	٠,٤٠	٠,٢٤
٦	٤٧	٠,٩٤	٠,٠٦	٠,٠٦

مجـ ب × ك = ١,١٠

٦ - نقوم بحساب التباين ($ع^{٢}$) لدرجات الاختبار فنجد أنه يساوى ٢,٨.

obeikandl.com

٧ - مما سبق يتضح أن :

$ع^{٢} = ٢,٨$

مجـ ب ك = ١,١٠

ن = ٦

وبتطبيق المعادلة K-R20 معادلة : (٢ - ٦) ينتج أن:

$$ر_{١٠١} = \frac{٦}{٥} \left(\frac{٢,٨ - ١,١}{٢,٨} \right)$$

$$= ١,٢ \left(\frac{١,٧}{٢,٨} \right)$$

$$= ١,٢ \times ٠,٦٠٧$$

$$= ٠,٧٣$$

إذا معامل الثبات المقدر للاختبار باستخدام معادلة K-R20 = ٠,٧٣

مثـــال : (٢ - ٥)

طبق اختـبار معرفـى يتكون من ١٥ فقـرة يجاب عليهـا بصح (✓) أو خطأ (×) على عدد ١٠ مفحوصين فكانت النتائج كالتالى:

مجـ س = ٨٠

سَ = ٨

مجـ س٢ = ٧٩٦

مجـ ب ك = ٢,٦٢

والمطلوب :

حساب معامل ثبات هذا الاختبار باستخدام معادلة K-R20.

الحـل :

١ - نقوم أولا بحساب التباين الكلى لدرجات الاختبار بتطبيق المعادلة (٢-٤) كالتالى :

$$ع_{ت} = \frac{١}{١٠} \sqrt{١٠(٧٩٦) - (٨٠)^{٢}}$$

$$= \frac{١}{١٠} \sqrt{٧٩٦٠ - ٦٤٠٠}$$

$$= \frac{١}{١٠} \sqrt{١٥٦٠}$$

$$= \frac{١}{١٠} \times ٣٩,٤٩٧$$

$$= ٣,٩٥$$

إذا $ع_{ت}^{٢} = (٣,٩٥)^{٢} = ١٥,٦٠$

٢ - نقوم بالتعويض فى المعادلة K- R20 كالتالى :

$$ر_{١٠١} = \frac{١٥}{١٥ - ١} \left(\frac{١٥,٦٠ - ٢,٦٢}{١٥,٦٠} \right)$$

$$= \frac{١٥}{١٤} \left(\frac{١٢,٩٨}{١٥,٦٠} \right)$$

$$= ١,٠٧١ \times ٠,٧٣٢$$

$$= ٠,٨٩١$$

إذا معامل الثبات المقدر للاختبار باستخدام معادلة K-R20 = ٠,٨٩

المعادلة K-R21 :

هذه المعادلة لا تتطلب تحليل الفقرات (العبارات أو الأسئلة) كما فى المعادلة السابقة، ولكنها تتطلب المتوسط الحسابى ($\bar{X}$ - سَ) لكل الاختبار وكذا الانحراف المعيارى (ع - SD) ، والصورة الرياضية لهذه المعادلة كالتالى :

obeikandl.com

$$\overline{\text{ر}}_{\text{ن ن}} = \frac{\text{ن ع}^2 - \overline{\text{س}}\ (\text{ن} - \overline{\text{س}})}{\text{ع}^2\ (\text{ن} - 1)}$$ معادلة (٢-٧)

حيث إن :

$\overline{\text{س}}$ = متوسط درجات الاختبار

ن = عدد وحدات (بنود) الاختبار

ع^2 = التباين الكلى لدرجات الاختبار

وبتطبيق هذه المعادلة على البيانات المعطاة فى المثال السابق (٢-٥) ينتج أن:

$$\overline{\text{ر}}_{\text{ن ن}} = \frac{١٥ \times ١٥,٦ - ٨\ (١٥ - ٨)}{٢,٨ \times (١٥ - ١)}$$

$$= \frac{٢٣٤ - ٥٦}{١٥,٦٠ \times ١٤}$$

$$= \frac{١٧٨}{٢١٨,٤}$$

$$= ٠,٨١٥$$

إذاً معامل الثبات المقدر للاختبار باستخدام معادلة K-R21 = ٠,٨٢ .

مثـال (٢-٦) :

لنفرض أننا طبقنا اختبار معرفى لكرة القدم يتكون من ١٠٠ عبارة (سؤال) على مجموعة من المدربين فكانت بياناتهم كالتالى:

$\overline{\text{س}}$ = ٦٥ درجة (متوسط درجات الاختبار)

ع = ١٠ درجات (الانحراف المعيارى)

والمطلوب :

حساب معامل ثبات هذا الاختبار باستخدام معادلة K-R21 .

obeikandl.com

الحل:

نقوم بالتعويض فى المعادلة (٢-٧) كالتالى:

$$\widehat{١٠١} = \frac{١٠٠ (١٠)^{٢} - ٦٥ (١٠٠ - ٦٥)}{(١٠)^{٢} (١٠٠ - ١)}$$

$$= \frac{١٠٠ \times ١٠٠ - ٦٥ \times ٣٥}{١٠٠ \times ٩٩}$$

$$= \frac{١٠٠٠٠ - ٢٢٧٥}{٩٩٠٠}$$

$$= \frac{٧٧٢٥}{٩٩٠٠}$$

$$= ٠,٧٨$$

معامل ألفا α :

فى عام ١٩٥١م قدم كرونباخ Cronbach, L.J معادلة عامة يمكن تطبيقها فى الحالات التى تكون فيها الاستجابة على الفقرة (العبارة أو السؤال) متعددة الاختيار، أى فى الحالات التى يكون للفقرة (العبارة أو السؤال) الواحدة عدد كبير من الدرجات المحتملة، حيث تكون احتمالات الإجابة ليست (صفرا) أو (واحداً) كما فى حالة الفقرات التى تتطلب الاستجابة بصح أو خطأ.

فمن المعروف أنه يوجد العديد من الاختبارات والمقاييس التى لا تتطلب الاستجابة عليها بصح أو خطأ كما فى مقاييس الشخصية والاتجاهات ، ففى مقياس كينيون Kenyon للاتجاهات نحو النشاط البدنى - على سبيل المثال - يقوم كل مفحوص بالإجابة عن كل عبارة بما يتناسب مع اتجاهه نحوها طبقا لمقياس مدرج من ٥ تدريجات هى: أوافق بدرجة كبيرة، أوافق، لم أكون رأيًا بعد، أعارض، أعارض بدرجة كبيرة. حيث لا تكون هناك إجابة صحيحة وأخرى خاطئة، فالإجابة التى يختارها المفحوص تعد مؤشراً عن اتجاهه بين الموافقة وعدم الموافقة بالنسبة لموضوع ما، وبحيث تمنح

الدرجـات: ٥، ٤، ٣، ٢، ١ أو ١، ٢، ٣، ٤، ٥ للتدريجـات الخمـسة بالترتيب حسب اتجاه العبارة.

ولكى يمكن تطبيـق معادلة كيودر - ريتشاردسون علـى مثل هذه المقاييس أعد كرونبـاخ معادلته المعروفة بمـعامل ألفا (α) coefficient alpha لتقدير الاتساق الداخـلى للاختبارات والمـقاييس متعـددة الاختيار، أى عنـدما تكون احتمالات الإجابة ليست صفراً أى ليست ثنائية البعد.

والصورة الرياضية لمعامل ألفا (α) كالتالى:

$$= \left(\frac{ن}{ن-١}\right)\left(١ - \frac{مجـ ع^{٢} ك}{ع^{٢}}\right)$$ معادلة (٢ - ٨)

حيث إن:

* α = معامل ألفا (معامل الثبات المقدر).

ن = عدد الفقرات.

ع٢ = تباين الدرجات على كل الاختبار.

مجـ ع٢ك = مجـموع تباين كل فقـرة من فقرات الاختـبار (من درجات المفحوصين على الفقرة).

طريقة تحليل التباين: ANOVA method

يطبق تحليل الـتباين analysis of variance (ANOVA) فى بعض الأحيان كمقيـاس لحساب ما يعرف بالثبـات النسبى relative reliability حيث يرجع الفضل فى استخدامه كأسلوب لتقدير ثبات الاختبارات والمقاييس إلى كل من جاكـسون Jackson ١٩٣٩م، وهويت Hoyt ١٩٤١م، والكسندر Alexander ١٩٤٧م. وكانت مجهودات هؤلاء العلماء قد بنيت على مسلمة رئيسية هى :

- أن الدرجة الكليـة للمفحوص على الاختـبار يمكن تقسيمـها إلى أربعة مكونات مـستقـلة هى: (أ) المكون الشـائع بين كل المـفحوصـين وكل فقرات الاختبار (ب) المكون المرتبط بفـقرات الاختبار فقط (جـ) المكون

المرتبط بالمفحوصين فقط د) مكون الخطأ وهو يرتبط بالمكونات الثلاثة السابقة. وحتى يمكننا فهم ما سبق فإنه يلزم توضيح الآتى:

١ - التباين الكلى (s^2t) : total variance

وهو يشير إلى مجموع مربعات انحرافات جميع القيم عن متوسطها الحسابى. ومن الملاحظ أن التباين الكلى لمجموعة من القياسات يتألف من مصدرين من مصادر التباين هما : التباين الحقيقى true variane وتباين الخطأ error variance، حيث يتم التعبير عن هذه العلاقة كالتالى:

الدرجة الملاحظة (التجريبية) = الدرجة الحقيقية + درجة الخطأ

$$س_م = س_ح + س_خ$$

حيث إن:

$س_م$ = الدرجة التى يتم الحصول عليها من التجربة نتيجة عملية القياس.

$س_ح$ = الدرجة الحقيقية.

$س_خ$ = درجة الخطأ (مكون الخطأ).

الدرجة الملاحظة (التجريبية) observed or experimental score

هى الدرجة التى يحصل عليها الفرد نتيجة الأداء على الاختبار، وتمتاز بأنها درجة يمكن ملاحظتها وتسجيلها (أى أنها درجة معلومة) تعبر عن الأداء العام أو الكلى للفرد.

الدرجة الحقيقة true score

وهى درجة تعبر عن الأداء الحقيقى للفرد، وتمثل جزءاً من الأداء العام أو الكلى على الاختبار (وهى درجة غير معلومة).

درجة الخطأ error score

هى درجة تعزى لأخطاء الصدفة أو للظروف الخارجية البعيدة عن موضوع الاختبار (وهى أيضا درجة غير معلومة).

أى أن:

التباين الكلى = التباين الحقيقى + تباين الخطأ

$$ع^{٢}_{م} = ع^{٢}_{ح} + ع^{٢}_{خ}$$ معادلة (٢ -٩)

ومن هذا يمكننا أن نقول إن :

$$\text{معامل الثبات} = \frac{\text{التباين الحقيقى}}{\text{التباين الكلى}}$$

$$= \frac{ع^{٢}_{ح}}{ع^{٢}_{م}}$$ معادلة (٢ - ١٠)

أو :

$$\text{معامل الثبات} = \frac{\text{التباين الحقيقى}}{\text{التباين الحقيقى + تبادين الخطأ}}$$

$$= \frac{ع^{٢}_{ح}}{ع^{٢}_{ح} + ع^{٢}_{خ}}$$ معادلة (٢ - ١١)

أى أن معامل ثبات درجات الاختبار تساوى النسبة بين التباين الحقيقى إلى التباين العام.

مثال : (٢-٧)

لنفرض أن لدينا مجموعة تتكون من ١٠ قياسات، وكان المتوسط الحسابى للدرجات الحقيقية هو ٢٥، وكان التباين الكلى للدرجات $ع^{٢}_{م}$ = ١٢٠، وتباين الدرجات الحقيقية $ع^{٢}_{ح}$= ١٠٥ ، وتباين الخطأ $ع^{٢}_{خ}$ = ١٥,٢ .

(جدول : ٢ - ٣) التالى :

جدول (٢ - ٣)

الدرجات الملاحظة	الدرجات الحقيقية	مكونات الخطأ
٣	٥	-٢
١٧	١٥	+٢
١٦	٢٠	-٤
٢٣	٢٥	-٢
٢٧	٢٥	+٢
٢٥	٢٥	٠
٣٥	٢٥	+١٠
٢٦	٣٠	-٤
٣٣	٣٥	-٢
٤٥	٤٥	٠
مجـ س = ٢٥٠	٢٥٠	٠
سَ = ٢٥	٢٥	٠
مجـ س٢ = ١٢٠٢	١٠٥٠	١٥٢
ع٢ = ١٢٠,٢	١٠٥	١٥,٢

obeikandl.com

من الجدول السابق يلاحظ الآتى:

بالنسبة للدرجات الملاحظة (التجريبية):

مجموع القيم مجـ س = ٢٥٠

المتوسط الحسابى سَ = ٢٥

مجموع مربعات انحرافات القيم عن وسطها الحسابى مجـ س٢ = ١٢٠٢

تباين الدرجات الملاحظة (التباين الكلى) ع٢م = ١٢٠,٢

وبالنسبة للدرجات الحقيقية :

مجموع القيم مجـ س = ٢٥٠

المتوسط الحسابى سَ = ٢٥

مجموع مربعات انحرافات القيم عن وسطها الحسابى مجـ س٢ = ١٠٥٠

تباين الدرجات الحقيقية ع٢ح = ١٠٥

وبالنسبة لتباين الخطأ :

مجموع القيم مجـ سَ = ٠

س = ٠

مجـ س٢ = ١٥٢

ع٢خ = ١٥,٢

وبتطبيق المعـادلة (٢ - ٩) يتضح أن التباين الكلى يسـاوى تباين الدرجات الحقيقية + تباين الخطأ :

١٢٠,٢ = ١٠٥ + ١٥,٢

١٢٠,٢ = ١٢٠,٢

ولقد استخدمت المعادلة (٢-٩) للـتأكد من أن التباين الكلى ع٢م الذى تم

حسابه من مجموع مربعات انحرافات القيم عن وسطها الحسابى يساوى تباين الدرجات الحقيقية $ع^{٢}_{ح}$ + تباين الخطأ $ع^{٢}_{خ}$.

ولحساب نسبة التباين الكلى إلى التباين الحقيقى نقوم بتطبيق المعادلة (٢-١٢) التالية :

$$\frac{ع^{٢}_{م}}{ع^{٢}_{م}} = \frac{ع^{٢}_{ح}}{ع^{٢}_{م}} + \frac{ع^{٢}_{خ}}{ع^{٢}_{م}}$$ معادلة (٢ - ١٢)

وبالتعويض فى المعادلة السابقة (٢-١٢) يتضح الآتى :

$$\frac{١٢٠,٢}{١٢٠,٢} = \frac{١٠٥}{١٢٠,٢} + \frac{١٥,٢}{١٢٠,٢}$$

$$١ = ٠,٨٧٤ + ٠,١٢٦$$

$$١ = ١$$

ويمكن حساب ثبات مجموعة الدرجات الموضحة بالجدول (٢ - ٣) بحساب نسبة تباين الدرجات الحقيقية $ع^{٢}_{ح}$ مقسوماً على التباين الكلى $ع^{٢}_{م}$ معادلة (٢ - ١٠) كالتالى :

$$\text{الثبات} = \frac{١٠٥}{١٢٠,٢} = ٠,٨٧$$

كما يمكن حساب الثبات بتطبيق المعادلة التالية (٢ - ١٣) كالتالى :

$$\text{الثبات} = ١ - \frac{١٥,٢}{١٢٠,٢}$$ معادلة (٢ - ١٣)

$$= ١ - \frac{ع^{٢}_{خ}}{ع^{٢}_{م}}$$

$$= ١ - ٠,١٢٦$$

$$= ٠,٨٧$$

obeikandl.com

ما هى القيم المقبولة لمعاملات الثبات؟

يتساءل معظم الباحثين عن قيم معاملات الثبات التى تجعلهم يثقون فى الأدوات التى سوف يستخدمونها فى جمع المادة العلمية فى الدراسات والبحوث التى يقومون بها، وتتوقف الإجابة على هذا التساؤل على الغرض من استخدام أداة القياس، فكثير من الباحثين يحكمون على مستوى معامل الثبات المحسوب من خلال الدلالة الإحصائية لمعامل الارتباط الذى يدل معامل الثبات، فإذا كان معامل الارتباط هذا دالاً إحصائيا كان معامل الثبات هو الآخر دالاً إحصائياً (أى مقبولاً) والعكس صحيح، إلا أن علماء القياس يرفضون هذا الأسلوب، ويؤكدون على أن القيمة المقبولة لمعامل الثبات المحسوب تتوقف أساساً على الغرض من استخدام أداة القياس (المقياس أو الاختبار)، ويتفقون فيما بينهم على ضرورة أن تكون هذه القيم كالتالى :

١- **بالنسبة للاختبارات التشخيصية** diagnosis واختبارات الذكاء Intelligence واختبارات الاستعدادات aptitudes واختبارات التحصيل achievement يجب ألا تقل قيم معاملات الثبات عن ٩٠,٠ .

٢- **فى الاختبارات الإكلينيكية** clinical يجب أن تكون معاملات الثبات مرتفعة جداً (تقترب من ٠٠,١).

٣- **فى الاختبارات التى تستخدم فى اتخاذ قرارات مهمة** عن مستقبل بعض الأفراد أو المؤسسات فإنه يلزم أن تكون معاملات الثبات ٩٥,٠ فأكثر.

٤- **فى الاختبارات التى تستخدم لمقارنة المتوسطات الحسابية** لمجموعتين أو أكثر من الأفراد فإنه يمكن قبول معامل الثبات الذى يساوى ٦٠,٠ كحد أدنى.

٤ - مفهوم الثبات المطلق وطرق حسابه

هو ثبات يعكس اتساق درجات الفرد على أساس الدرجات الملاحظة (التجريبية)، حيث يستخدم الخطأ المعيارى للقياس standard error of measurement كمؤشر لوصف ثبات الاختبار وإقرار مدى كفاية وصلاحية هذا الثبات. فالخطأ المعيارى طريقة إحصائية تستخدم لإيضاح مفهوم الثبات وملاحظة التغيرات (الاختلافات) التى تحدث فى درجات الاختبار عند تطبيقه لعدد من المرات على نفس الأفراد مع توحيد الشروط فى جميع المرات.

ويستخدم الخطأ المعيارى (SE) فى القياس لتحديد مدى الاختلاف الحادث فى درجات المقياس أو الاختبار عند تطبيقه لعدد من المرات على نفس الأفراد، فإذا طبق الاختبار ثم أعيد تطبيقه أى عدد من المرات على نفس الأفراد، مع افتراض عدم حدوث تغير فى الصفة أو القدرة أو السمة المقيسة، فإن الانحراف المعيارى(ع - SD) لتوزيع الدرجات الفعلية فى كل مرة من مرات القياس يشير إلى الخطأ المعيارى (SE) الذى يحدث فى عملية القياس أو فى درجات الاختبار الذى طبق على الأفراد.

فمعامل الثبات يدل على اتساق درجات الاختبار عندما يتم تطبيقه أكثر من مرة على نفس الأفراد، حيث يكون هذا المعامل منخفضًا فى حالة وجود اختلافات كبيرة فى الدرجات، ويكون مرتفعًا عندما يكون الاختلاف فى الدرجات محدودًا.

ويستهدف الثبات المطلق تطبيق أداة القياس عدد من المرات على نفس الفرد (المفحوص) حتى يمكن تقدير كمية التغير variation التى يمكن توقعها بالنسبة لدرجات الاختبار، هذا التقدير estimate يسمى بالخطأ المعيارى للقياس (SE).

ومن الملاحظ أن دليل الاختبارات الكتابية المقننة (الاختبارات المنشورة) يتضمن - عادة - الخطأ المعيارى للقياس كجزء مهم من شروط الاختبار، وفى مجال الاختبارات النفس - حركية يلعب الخطأ المعيارى للقياس دوراً مهمًا بالنسبة لتقدير درجة المفحوص على هذه الاختبارات.

ما الذى يعنيه الخطأ المعيارى للقياس؟

لنفترض أنه تم تطبيق اختبار الرمية الحرة فى كرة السلة free throw ٢٠ مرة على ١٠٠ لاعب، وقد حصل أحد اللاعبين على الدرجة ١٢ من المحاولات العشرين، وكان الخطأ المعيارى للقياس (SE) يساوى ٢، فما الذى يعنيه هذا الرقم بالرجوع إلى درجة هذا اللاعب والتى تساوى ١٢؟

إن هذه الدرجة تظهر كمية الخطأ amount of error التى يمكن وضعها فى الاعتبار عند تحليل الدرجة التى حصل عليها اللاعب والتى تساوى ١٢. بمعنى آخر، تدلنا الدرجة ٢ على الحدود limits التى تمتد إليها الدرجة الحقيقية للقدرة على الرمية الحرة لهذه اللاعب، فإذا كانت درجات اختبار الرمية الحرة ثابتة تماماً perfect reliabitity ، فإنه فى حالة تكرار تطبيق الاختبار على هذا اللاعب أى عدد من المرات تحت نفس الشروط، فإن حوالى ٦٨٪ من الدرجات الحقيقية سوف تقع فى مدى وحدة خطأ معيارى واحدة بالنسبة للدرجة الملاحظة (التجريبية) observed score (١٢ ± ٢) ، إن حوالى ٩٥٪ من الدرجات سوف تقع فى مدى وحدتين خطأ معيارى (١٢ ± ٤)، وأن حوالى ٩٩٪ من الدرجات سوف تقع فى مدى ثلاث وحدات خطأ معيارى (١٢ ± ٦).

وقد أخذ بهذه القاعدة وفقًا للمساحات تحت منحنى التوزيع الاحتمالى، وعلى أساس أن الدرجة الملاحظة (التجريبية) تم تحويلها إلى الدرجة ذ - z ، وأن الخطأ المعيارى (SE) يكافئ درجة واحدة من الدرجات ذ - z.

ووفقاً لهذه القاعدة فإنه يمكن حساب مستوى الثقة فى درجات الرمية الحرة للاعب فى المثال السابق على النحو التالى :

عدد وحدات الخطأ المعيارى (SE)	مستوى الثقة (L.C)	وحدات الخطأ للاختبار	مدى الدرجات (R.S)
١	٦٨٪	٢	١٠ - ١٤
٢	٩٥٪	٤	٨ - ١٦
٣	٩٩٪	٦	٦ - ١٨

وعلى هذا الأساس نستطيع أن نقرر أن الدرجة الحقيقية لهذا اللاعب فى اختبار الرمية الحرة فى كرة السلة تنحصر ما بين (٦ إلى ١٨ درجة) خلال ثلاث وحدات انحراف معيارى ومستوى ثقة ٩٩٪، فى حين أن درجة الأداء الفعلية لهذا اللاعب بالنسبة لنفس الاختبار تميل إلى أن تنحصر ما بين (١٠ إلى ١٢ درجة) .

ويمكن حساب الخطأ المعيارى للقياس (SE) باستخدام المعادلة التالية :

$$ع_{ت} = ع \sqrt{1 - ر_{١٠١}}$$

حيث إن :

$ع_{ت}$ = الخطأ المعيارى لعملية القياس (درجات الاختبار).

ع = الانحراف المعيارى لدرجات الاختبار.

$ر_{١٠١}$ = معامل ثبات درجات الاختبار.

وتستخدم هذه المعادلة لتقدير الخطأ المعيارى (SE) لعملية القياس التى تنتج من تكرار تطبيق الاختبار (القياس)، حيث يدل الخطأ المعيارى على مدى تطابق الدرجات الفعلية التجريبية Actual scores بالنسبة للدرجات الحقيقية True scores عند تطبيق الاختبار عدد من المرات.

فإذا افترضنا أن الانحراف المعيارى(ع – SD) فى المثال السابق الخاص باختبار الرمية الحرة فى كرة السلة كان يساوى ٤، وكان معامل الثبات يساوى ٠,٧٥ فإنه يمكن حساب الخطأ المعيارى للقياس بالتعويض فى معادلة الخطأ المعيارى كالتالى :

$$ع_{ت} = ٤ \sqrt{١ - ٠,٧٥}$$

$$= ٤ \sqrt{٠,٢٥}$$

$$= ٤ \times ٠,٥$$

$$= ٢$$

إذا الخطأ المعيارى للقياس يساوى فى هذه الحالة ٢ .

وإذا افترضنا أن الخطأ المعيارى للقياس كان يساوى ٠,٣٣٣ بدلاً من ٢، فإن مستوى الثقة فى درجات اختبار الرمية الحرة فى المثال السابق يصبح كالتالى:

(١٢ ± ٠,٣٣٣) = ١١,٦٦٧ إلى ١٢,٣٣٣ عند مستوى ثقة ٦٨٪.

(١٢ ± ٠,٦٦٦) = ١١,٣٣٤ إلى ١٢,٦٦٦ عند مستوى ثقة ٩٥٪.

(١٢ ± ٠,٩٩٩) = ١١,٠٠ إلى ١٣,٠٠ عند مستوى ثقة ٩٩٪.

حيث يلاحظ أن الدرجة تنحصر ما بين ١١و ١٣ درجة بدلاً من ٦و ١٨ درجة.

العلاقة بين الانحراف المعيارى والخطأ المعيارى:

يزداد معامل ثبات الاختبار كلما نقص الانحراف المعيارى (ع – SD) والخطأ المعيارى ($ع_{ن}$ – SE)، لهذا السبب كان لمعاملات الثبات المرتفعة أخطأ صغيرة، فى حين تتميز معاملات الثبات المنخفضة بوجود أخطاء كبيرة. فإذا كان معامل ثبات أحد الاختبارات يساوى ٠,٨٤، وكان الانحراف المعيارى لدرجات الاختبار يساوى ١٠ ، فإن الخطأ المعيارى للقياس يمكن حسابه كالتالى :

$$ع_{ن} = ١٠ \sqrt{١ - ٠,٨٤}$$

$$= ١٠ \sqrt{٠,١٦}$$

$$= ١٠ \times ٠,٤$$

$$= ٤,٠٠$$

وإذا كان الانحراف المعيارى يساوى ٥ بدلاً من ١٠ ، فإن الخطأ المعيارى لعملية القياس (درجات الاختبار) يكون كالتالى:

$$ع_{ت} = ٥ \sqrt{١ - ,٨٤}$$

$$= ٥ \times ,٤$$

$$= ٢,٠٠$$

ويلاحظ أن الخطأ المعيارى للقياس انـخفض إلى النصف ، عندما انخفض الانحراف المعيـارى للدرجات، وذلك على أساس افتراض أن معـامل الثبات ظل ثابتا.

مؤشر (دليل) الثبات : index of reliabitity

هو مفـهوم نظرى يشير إلـى العلاقة بين الدرجـات الملاحظة (التجريـبية) والدرجات الحقـيقية التى خلصت مـن أخطأ القياس. ويمكن حـساب مؤشر (دليل) الثبات باستخدام المعادلة التالية:

مؤشر الثبات = $\sqrt{ر_{١٠١}}$

حيث إن:

$ر_{١٠١}$ = معامل ثبات الاختبار (الدرجات الملاحظة أو التجريبية).

٥ - الثبات الخارجى والثبات الداخلى

هناك وجهة نظر ترى أنه يمكن تصنيف الثبات إلى نمطين عريضين هما :

الثبات الخارجى : interclass reliabitity

وهو يستخـدم عندما يتم أخذ القيـاسات لمرتين فقط، حيـث يتم حساب المعاملات الخارجية interclass coeffients باستخدام معامـل الارتباط التتابعى (ppm) بطريقة بيرسون، ويتضمن هذا النمط الطرق الآتية :

١ - ثبات إعادة الاختبار (استقرار درجات الاختبار)

test-retest reliability

٢ - ثبات الصور المتكافئة (المتبادلة - المتوازية - المتساوية)
equivalent reliability

٣ - ثبات نصف الاختبار split - halves reliability

حيث يستخدم لتصحيح ثبات نصف الاختبارات المعادلات التالية :

- معادلة سبيرمان - براون.
- معادلة رولون.
- معادلة فلانجان.

الثبات الداخلى : Intraclass reliability

وهو نمط من الثبات يختلف عن سابقة فهو يستخدم عندما يتم إجراء القياسات لثلاثة محاولات متتالية أو أكثر، حيث يستخدم تحليل التباين ANOVA كأسلوب إحصائى لتقدير الثبات. ويتضمن الثبات الداخلى استخدام بعض الأساليب الإحصائية المهمة هى :

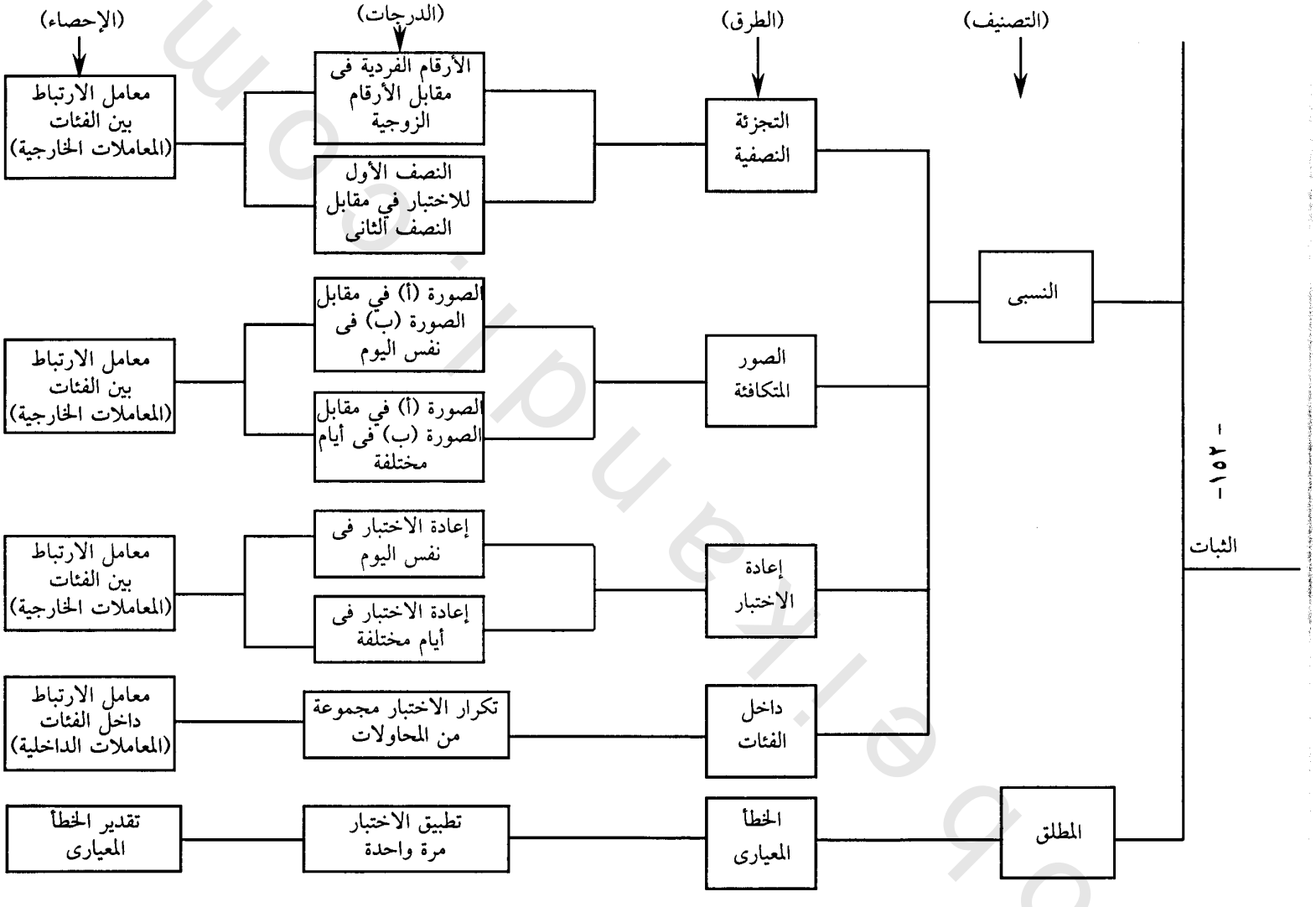
١ - معادلات كيودر - ريتشاردسون (KR21) (KR20).

٢ - معامل ألفا ∝ Coeffient alpha.

٣ - تحليل التباين (ANOVA).

خريطة الثبات : Reliability chart

يعرض شكل (٢-١) ملخصًا لتصنيف الثبات، والطرق المستخدمة لحساب كل تصنيف وفقًا لطبيعة عملية القياس من حيث عدد مرات التطبيق، والفترات الزمنية بين مرات التطبيق المختلفة، هذا بالإضافة إلى الأساليب الإحصائية المناسبة لتقدير معاملات الثبات بالنسبة لكل طريقة على حدة.



شكل (٢ - ١) خريطة الثبات

٦- العوامل التى تؤثر على ثبات درجات الاختبار

يتأثر ثبات درجات الاختبار بعدد كبير من العوامل، هذه العوامل يمكن وضعها فى فئتين عريضتين هما:

- عوامل خارجية Ixtrinsic factors
- عوامل داخلية Intrinsic factors

العوامل الخارجية:

وهى عوامل تحدث بعيداً عن الاختبار نفسه وتؤدى إلى أن يصبح الاختبار ثابتًا أو غير ثابت، ولعل من أهم هذه العوامل ما يأتى:

١) تشتت مجموعة الأفراد : Group Variability

فعندما تكون مجموعة المفحوصين متجانسة homogeneous فى القدرة المزمع قياسها، فمن المرجح الحصول على معاملات ثبات منخفضة، أما إذا كانت مجموعة المفحوصين تتفاوت فيما بينها تفاوتاً كبيراً بالنسبة للقدرة المقيسة، فمن المرجح الحصول على معاملات ثبات مرتفعة.

٢) التخمين: Guessing

ويقصد به تخمين المفحوصين بالنسبة للعبارات (الأسئلة) الخاصة باختبارات الورقة والقلم، والذى يعد من أهم مصادر عدم الثبات، ففى حالة الاستجابات ثنائية البعد (صح أو خطأ) تكون فرصة الإجابات الصحيحة على العبارات على أساس التخمين هى ٥٠٪، فى حين ينخفض تأثير التخمين بالنسبة للعبارات التى تتطلب الاستجابة عليها الاختيار من متعدد (Mcq).

وبشكل عام فقد يؤدى التخمين إلى زيادة الدرجة الكلية للاختبار فيزيد بذلك معامل الثبات، أو يزيد من أخطاء القياس مما يؤدى إلى انخفاض معامل الثبات.

٣) **الاشتراطات البيئية** : Environmental Conditions

وتتمثل هذه الاشتراطات فى مدى ضبط وتنظيم بيئة الاختبار مثل: الإضاءة، الضوضاء، درجات الحرارة، واتجاهات الرياح، ونسبة الرطوبة، والزى (وبخاصة فى الاختبارات البدنية) إلى غير ذلك من العوامل التى يكون لها تأثير على ثبات درجات الاختبار.

٤) **الأحداث الطارئة** : Momentary Fluctuations

تؤثر بعض الأحداث الطارئة على ثبات درجات الاختبار إما بالزيادة أو بالنقصان، فالأحداث الطارئة التى تحدث للمفحوص أثناء الأداء على الاختبار قد تؤدى إلى الاضطراب والتشويش وعدم التركيز، فالأصوات العنيفة المفاجئة، وتداعى بعض المواقف والأحداث غير السارة، وضياع القلم أو كسره قبل أو أثناء الإجابة، جميعها عوامل قد تؤدى إلى قلق وتوتر المفحوص فيقع فى أخطاء تؤثر فى ثبات درجات الاختبار.

العوامل الداخلية:

وهى عوامل تختص بالاختبار نفسه وتؤثر على ثباته، ومن أهم هذه العوامل ما يأتى:

١) **طول الاختبار** : length of the test

يقصد بطول الاختبار عدد الوحدات (البنود - الفقرات - الأسئلة) التى يتضمنها، فكلما زاد طول الاختبار زاد بالتالى ثبات هذا الاختبار، فالاختبارات الأطول تميل إلى إظهار معاملات ثبات أعلى من الاختبارات الأقصر. وتتأسس هذه النتيجة على حقيقة علمية مؤداها أن الاختبارات الطويلة تتأثر بدرجة أقل بعوامل الصدفة والتخمين من الاختبارات القصيرة.

ولتوضيح ذلك، لنفترض أننا أردنا قياس القدرة على التهجئة Spelling ability فقمنا بسؤال التلاميذ لتهجئة كلمة واحدة فقط، فى هذه الحالة سوف نحصل على نتائج غير ثاتبة، حيث يرجع ذلك إلى أن التلاميذ الذين استطاعوا الإجابة على الكلمة نجحوا، والتلاميذ الذين لم يستطيعوا الإجابة على نفس الكلمة رسبوا فى الاختبار. فإذا كانت الكلمة «صعبة» فإن معظم التلاميذ سوف يرسبون، وإذا كانت «سهلة» فإن معظم التلاميذ سوف ينجحون، وعليه فإن كلمة واحدة (عبارة أو فقرة أو سؤال) سوف يظهر عدم ثبات درجات الاختبار عندما يستخدم هذا الاختبار لتقدير قدرة التلميذ على التهجئة.

فإذا ضمنا اختبار التهجئة السابق قائمة من الكلمات المتساوية فى درجة صعوبتها، حينئذ يمكننا تقدير قدرة التلميذ على التهجئة بطريقة جيدة، وذلك لكون الدرجات التى يتم الحصول عليها من عدد كبير من الكلمات تظهر بشكل أفضل الفروق الحقيقية فى القدرة على التهجئة ومن ثم تصبح أكثر استقراراً فعن طريق زيادة حجم عينة الأسئلة (الفقرات أو البنود) التى يتضمنها الاختبار يزيد بالتالى الاتساق الداخلى لهذا الاختبار.

مما سبق يتضح أن طول الاختبار يؤثر على ثباته، حيث يتحدد هذا التأثير **من خلال نقطتين مهمتين هما:**

أ - أن العلاقة بين طول الاختبار (عدد الوحدات) وثباته علاقة طردية.

ب - أنه كلما زاد عدد الوحدات (الأسئلة أو البنود) ارتفع معامل ثبات الاختبار. ويستخدم لتقدير معامل ثبات الاختبار بعد زيادة طوله معادلة سبيرمان - براون التالية :

$$\text{ر}_{1\cdot1} = \frac{\text{ن} \times \text{ر}}{1 + (\text{ن} - 1)\,\text{ر}} \qquad \text{معادلة (٢ - ١٤)}$$

حيث إن :

$\rho_{1\cdot1}$ = معامل ثبات الاختبار بعد زيادة طوله (عدد وحداته)

ρ = معامل ثبات الاختبار قبل زيادة طوله (عدد وحداته)

ن = عدد مرات زيادة الاختبار (نسبة عدد الوحدات بعد الزيادة إليها قبل الزيادة)

مثال : (٢ - ٨)

طبق اختبار للذكاء يتكون من ١٠٠ سؤال (عبارة أو فقرة)، وكان معامل ثبات هذا الاختبار ٠,٨٠، فإذا زاد هذا الاختبار ٤ مرات بالنسبة لطوله، فكم يكون معامل ثبات الاختبار بعد زيادة طوله؟

الحل :

$$\rho_{1\cdot1} = \frac{٤ \times ٠,٨٠}{١ + (٤ - ١) ٠,٨٠} \quad \text{معادلة (٢ - ١٤)}$$

$$= \frac{٣,٢}{١ + ٣ \times ٠,٨٠}$$

$$= \frac{٣,٢}{١ + ٢,٤}$$

$$= \frac{٣,٢}{٣,٤}$$

$$= ٠,٩٤$$

يلاحظ من المثال السابق أن زيادة طول الاختبار ٤ مرات بالنسبة لطوله الأصلي زاد من معامل ثبات هذا الاختبار - عندما طبق على نفس المفحوصين - من ٠,٨٠ إلى ٠,٩٤.

obeikandl.com

مثـــال : (٢ - ٩)

طبق اختـبار للذكاء يتكـون من ١٠٠ سؤال وكان معامـل ثباته ٠,٨٠ ، فكم يكون معامل ثبات هذا الاختبار إذا أصبح يتكون من ٤٠٠ سؤال ؟

الحـــل:

١ - نقوم بحسـاب (ن) والتى تمثل النسبة بين عدد الـوحدات بعد الزيادة إلى عدد الوحدات قبل الزيادة كالتالى :

$$\text{ن} = \frac{٤٠٠}{١٠٠} = ٤$$

٢ - نقوم بالتعويض فى معادلة سبيرمان - براون (معادلة : ١٤) كالتالى:

$$\text{١٠١} = \frac{٤ \times ٠,٨٠}{١ + (٤ - ١) ٠,٨٠}$$ معادلة (٢ - ١٤)

$$= \frac{٣,٢}{٣,٤}$$

= ٠,٩٤ وهى نفس النتيجة فى المثال السابق.

مثال : (٢ - ١٠)

طبق اختـبار معرفى للثـقافة الصحـية يتكون من ٥٠ سؤلاً (فـقرة أو بند) وكان معامل ثبات هذا الاختبار ٠,٧٥ فكم يكون معامل ثباته إذا أصبح عدد وحداته ١٥٠ سؤالاً ؟

١ - نقوم أولاً بحساب (ن) كالتالى:

$$\text{ن} = \frac{١٥٠}{٥٠} = ٣$$

٢ - نقوم بتطبيق (معادلة : ١٤) كالتالى:

obeikandl.com

$$\rho_{1\cdot1} = \frac{3 \times 0{,}75}{1 + (3-1)\ 0{,}75} \qquad \text{معادلة (٢ - ١٤)}$$

$$= \frac{2{,}25}{1+2\times0{,}75}$$

$$= \frac{2{,}25}{1 + 1{,}5} = \frac{2{,}25}{2{,}50}$$

$$= 0{,}90$$

مثال : (٢ - ١١)

طبق اختبار للثقافة الصحية معامل ثباته ٠,٧٥ عندما كان عدد وحداته ٥٠ سؤلاً (فقرة أو بند) ، فكم يصبح معامل ثبات هذا الاختبار إذا أضيف إلى عدد وحداته ١٠٠ سؤال؟

الحل:

١ - نقوم أولاً بحساب عدد الوحدات بعد الزيادة كالتالى:

عدد الوحدات بعد الزيادة = ٥٠ + ١٠٠ = ١٥٠

٢ - نقوم بحساب (ن) كالتالى:

$$\text{ن} = \frac{150}{50} = 3$$

obeikandl.com

٣ - نقوم بالتعويض فى المعادلة (٢ - ١٤) كالتالى :

$$\text{ر}_{\text{١٠١}} = \frac{٣ \times ٠,٧٥}{١ + (٣-١)\ ٠,٧٥}$$

$$= ٠,٩٠ \qquad \text{(وهى نفس النتيجة السابقة)}$$

مثال : (٢ - ١٢)

طبق اختبار للثقافة الصحية يتكون من ٥٠ سؤلاً ومعامل ثباته ٠,٧٥ ، والمطلوب أن يكون معامل ثبات هذا الاختبار ٠,٩٠ ، فكم يكون عدد وحداته؟

الحل :

بالتعويض فى المعادلة الأساسية (٢ - ١٤) ينتج أن :

$$\text{ر}_{\text{١٠١}} = \frac{\text{ن} \times \text{ر}}{١ + (\text{ن}-١)\ \text{ر}}$$

$$٠,٩٠ = \frac{\text{ن} \times ٠,٧٥}{١ + (\text{ن}-١)\ ٠,٧٥}$$

$$٠,٧٥\ \text{ن} \times ٠,٩٠ = ٠,٧٥ [١ + (\text{ن} - ١)$$

$$٠,٧٥\ \text{ن} ٠,٩٠ = ٠,٩٠ + (٠,٩٠\ \text{ن} - ٠,٩٠)٠,٧٥$$

$$٠,٧٥\ \text{ن} ٠,٩٠ = ٠,٩٠ + (٠,٦٨\ \text{ن} - ٠,٦٨)$$

$$٠,٧٥\ \text{ن} ٠,٩٠ = ٠,٢٢ + ٠,٦٨\ \text{ن}$$

$$٠,٧٥\ \text{ن} - ٠,٦٨\ \text{ن} = ٠,٢٢$$

$$٠,٠٧\ \text{ن} = ٠,٢٢$$

obeikandl.com

ن = $\frac{٢٢}{٧}$ = ٣,١٤ = ٣ مرات تقريبا حيث تستبعد الكسور الأقل من ٠,٥

مما سبق يتضح أنه لكى نرفع معامل ثبات اختبار الثقافة الصحية من ٠,٧٥ إلى ٠,٩٠ ، فإنه يجب أن نزيد عدد أسئلته من ٥٠ سؤلاً إلى ١٥٠ سؤلاً، حيث (ن = ٣) أى ٣ × ٥٠ = ١٥٠

ولتسهيل العمليات الحسابية فى المثال (٢ - ١٢) يمكن تطبيق المعادلة التالية لحساب قيمة (ن) بطريقة مباشرة، وهذه المعادلة هى :

$$\text{ن} = \frac{\text{معامل الثبات المطلوب} \times ١ - \text{المعامل الحالى}}{\text{المعامل الحالى} \times ١ - \text{المعامل المطلوب}}$$ معادلة (٢ - ١٥)

وبتطبيق هذه المعادلة على المعطيات فى المثال (٢ - ١٢) ينتج أن:

$$= \frac{٠,٩٠ \times (١ - ٠,٧٥)}{٠,٧٥ \times (١-٠,٩٠)}$$

$$= \frac{٠,٩٠ \times ٠,٢٥}{٠,٧٥ \times ٠,١}$$

$$= \frac{٠,٢٢٥}{٠,٧٥}$$

إذاً ن = ٣ (وهى نفس النتيجة السابقة تقريبا)

٢) درجة صعوبة الفقرات : Difficulty value of items

من المعروف أن درجة المفحوص على الاختبار أو المقياس هى مجموع أو محصلة درجاته على فقراته، فالفقرة هى الوحدة التى تكسب المقياس خصائصه، حيث تتوقف الخصائص الإحصائية للمقياس على المعالم الإحصائية لفقراته، فقد تبين أن عدد ومستوى صعوبة الفقرات يؤثر على ثبات وصدق المقياس أو الاختبار.

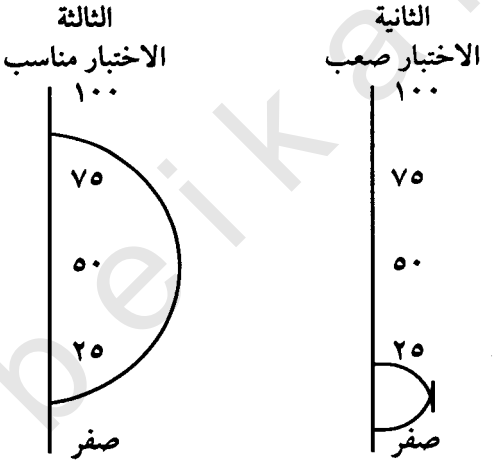
فالفقرات التى لـها مؤشر صعوبة يساوى ٥, ٠ يكون لها مـعامل ثبات أعلى، من الفقرات التى لهـا مؤشرات صعوبة أكبر من ٥, ٠ ، كما أن الـفقرات السهلة جداً والفقرات الـصعبة جداً تؤدى إلى معـاملات ثبات منخفضة، ويـبين الشكل (٢- ٢) مقارنة افتراضية لتوزيعات درجات اخـتبار للتحصيل الدراسى يتكون من ١٠٠ فقرة ومعاملات ثبات درجات هذا الاخـتبار عندما طبق على ثلاث عينات مختلفة من التلاميذ، حيث يتضح الآتى :

الحالة الأولى : عندمـا بلغ معامـل الصعوبـة ٢٥, ٠ ، نجد أن معامـل ثبات الاختبار بلغ ٥٠, ٠ .

الحالة الثـانية : عندمـا بلغ مـعامل الصعوبة ٧٥, ٠ نجد أن أن معـامل الثبات بلغ ٥٠, ٠

الحالة الثالثة : عندما بلغ معامل الصعوبة ٥, ٠ نجد أن معامل الثبات بلغ ٩٠, ٠

الحـــالات



	الأولى	الثانية	الثالثة
المتوسط الحسابى =	٨٨	١٣	٥٠
المدى =	٢٥	٢٥	١٠٠
الثبات =	٠,٥٠	٠,٥٠	٠,٩٠

شكل (٢ - ٢)

٣) مدى الدرجات : Range of scores

عندما يكون مدى الدرجات الكلية للاختبار ضيقا (محدوداً) يقل الانتشار (التشتت) Variability فينخفض بذلك معامل الثبات، ففى الشكل السابق (٢ - ٢) يظهر أن مدى الدرجات يساوى فى الحالة الأولى ٢٥ (١٠٠ - ٧٥=٢٥)، ويساوى فى الحالة الثالثة حاولى ١٠٠ (١٠٠ - ٠=١٠٠)، وعليه يظهر أنه كلما كان انتشار درجات الاختبار كبيرا كان معامل ثبات هذه الدرجات هو الآخر كبيراً، والعكس صحيح.

ومن ناحية يمكننا القول أنه كلما كان الانحراف المعيارى (ع -SD) للدرجات الكلية للاختبار مرتفعا كان معامل الثبات هو الآخر مرتفعا، وكلما كان الانحراف المعيارى منخفضا كان معامل الثبات هو الآخر منحفضا.

٤) تجانس الفقرات : Homogeneity of items

يعد تجانس الفقرات أحد أهم العوامل التى تؤثر على الثبات، حيث يقصد بالتجانس هنا الآتى:

- معامل ثبات الفقرة (معاملات ارتباط الفقرة مع غيرها)
- السمات والخصائص والقدرات التى تقيسها كل فقرة (وظائف الفقرة).

فعندما تقيس الفقرات وظائف مختلفة وتكون معاملات الارتباط البينية فيما بين الفقرات Intercorrelations صفرية (أو قريبة من الصفر) فإن معامل الثبات فى هذه الحالة يظهر منخفضًا جداً (صفر أو قريب منه) وذلك بسبب عدم تجانس الفقرات Hetrogeneous فى الوظائف التى تقيسها، أما إذا كانت الفقرات تقيس نفس الوظائف (نفس السمة أو القدرة) حينئذ تظهر معاملات الارتباط البينية (فيما بين الفقرات) مرتفعة، وبالتالى يظهر ثبات الاختبار هو الآخر مرتفعًا.

٥) **القيمة التميزية للفقرة** : Discriminative value

ويقصد بها الارتباط بين كل فقرة من فـقرات الاختبار والدرجة الكلية له، وتعرف هذه العـملية بتحليـل الفقرة item analysis والتى تبيـن الارتباط بين درجة الأداء على الفقرة والدرجة الكلية للاخـتبار. فعندما يكون الارتباط بين درجة الفقـرة والدرجة الكلية لـلاختبار منخـفضاً، فإن ذلك يعنـى أن الفقرة تقيس شيء آخر مختلف عما تقيسة الفقرات الأخرى للاختبار، كما قد يعنى هذا الارتباط أيضا أن الفقرة كانت سهلة جداً أو صعبة جداً حيث لم تقدم لنا فروقاً ظاهرة فيما بين استجابات المفحـوصين عليها، وعليه يظهر معامل ثبات الاختبار منخفضاً حيث يفضل فى مثل هذه الحالة حذف مثل هذه الفقرة، أما إذا كان الارتباط بين الفقرة والدرجة الكلية للاخـتبار مرتفعاً فإن معامل ثبات الاختبار يصبح هو الآخر مرتفعاً.

٧ - التصحيح للتنقية

تستخدم معاملات الثبات للاختبـارات والمقاييس المختلفة لتصحيح الارتباط بين أى مقياسـين انطلاقاً من أن الارتباط بين أى مقياسـين لا يخلو من الخطأ الذى حدث أثناء عمليات القياس.

وتستـخدم عملية التـصحيح للتنـقية Correction for Attenuation كطريقة لتصحيح معاملات الارتباط بين الاختبارات من أخطاء القياس Measurement errors ، حيث تستلزم هذه الطريقة توافر الآتى:

١ - معاملات ثبات الاختبارين (كل اختبار على حدة).

٢ - معامل الارتباط بين الاختبارين.

ويستخدم لذلك المعادلة التالية:

$$\text{ر} = \frac{\text{ر}_{٢٠١}}{\sqrt{\text{ر}_{١٠١} \times \text{ر}_{٢٠٢}}}$$ معادلة (٢ - ١٦)

حيث إن :

ر = الارتباط الحقيقى بين الاختبارين (١ ,٢)

$\text{ر}_{٢٠١}$ = الارتباط الملاحظ (التجريبى) بين الاختبارين (١ ,٢)

$\text{ر}_{١٠١}$ = معامل ثبات الاختبار الأول (١)

$\text{ر}_{٢٠٢}$ = معامل ثبات الاختبار الثانى (٢)

مثال : (٢ - ١٣)

حسب الارتباط بين اختبار للتوافق بين العين وحركة اليد واختبار للمهارة الرياضية العامة general athletic skill فكان هذا الارتباط ٠,٣٤ ، فإذا كان معامل ثبات الاختبار الأول ٠,٧٥ ، ومعامل ثبات الاختبار الثانى ٠,٨٢ ، فما هو الارتباط الحقيقى (الخالى من أخطاء القياس) بين الاختبارين.

الحل :

نقوم بتطبيق المعادلة (٢ - ١٦) كالتالى:

$$\text{ر} = \frac{٠,٣٤}{\sqrt{٠,٧٥ \times ٠,٨٢}}$$

$$= \frac{٠,٣٤}{\sqrt{٠,٦١٥}} = \frac{٠,٣٤}{٠,٧٨} = ٠,٤٤$$

يتبين من المثال السابق أن الارتباط بين الاختبارين قد زاد من ٠,٣٤ إلى ٠,٤٤ عندما استخدمت طريقة التصحيح.

obeikandl.com

كيف نحسن من ثبات درجات الاختبار؟

يمكن تحسين ثبات درجات الاختبار عن طريق ضبط جميع العوامل الخارجية والداخلية التى توثر على الثبات والتى سبق الإشارة إليها، ولضمان الحصول على معاملات ثبات جيدة فإنه يجب مراعاة الشروط التالية:

١ - يجب أن تكون عينة المفحوصين غير متجانسة بالنسبة للقدرة أو السمة أو الخاصية التى يقيسها الاختبار.

٢ - يجب أن تكون وحدات (الأسئلة أو الفقرات) متجانسة.

٣ - يفضل أن يكون الاختبار طويلا.

٤ - يفضل أن تكون وحدات الاختبار متوسطة الصعوبة، بمعنى أن تتراوح مؤشرات الصعوبة من ٠,٤٠ إلى ٠,٦٠

٥ - أن تتمتع الفقرات بالقدرة على التمييز.

تأثير العمر الزمنى على الثبات:

أظهرت التجارب العلمية المختلفة أن ثبات درجات الاختبار يزداد كلما تقدم السن (النضج)، وأنه توجد فروقاً بين البنين والبنات لصالح البنين، فقد أظهرت نتائج التجارب التى أجريت لحساب ثبات اختبار ستانفورد بينية Stanford- Binet scale لنسبة الذكاء 1Q عن طريق إعادة اختبار معاملات الثبات التالية:

معاملات الثبات

فئات السن	بنين	بنات
٦ - ١٠ سنوات	٠,٧٧	٠,٧٤
٧ - ١٠ سنوات	٠,٧٩	٠,٧٦
٨ - ١٠ سنوات	٠,٨٩	٠,٨٣
٩ - ١٤ سنة (للجنسين)	٠,٨٢	

كما أظهرت التجارب أن طريقة إعادة الاختبار هى أكثر طرق حساب الثبات التى تتأثر بالسن (النضج)، وأنه كلما تقدم السن مع توحيد الفترة الزمنية بين التطبيقين الأول والثانى زاد الثبات، فقد سجل فريمان Freeman ١٩٦٥م. لاختبار ستانفورد بينية معاملات الثبات التالية:

السن عند بدء الاختبار	معامل الثبات عن طريق إعادة الاختبار
٤ سنوات	٠,٧٣
٩ سنوات	٠,٨٧
١١سنة	٠,٩٢

ومن ناحية أخرى أظهرت التجارب العلمية أن طول الفترة الزمنية بين التطبيقين الأول والثانى يؤدى إلى انخفاض معاملات الثبات بالنسبة لاختبارات الذكاء والاستعدادات العقلية، فقد حصل «فريمان» ١٩٦٥م على معاملات الثبات التالية:

الفترة الزمنية	معامل الثبات
- إعادة الاختبار بعد يوم واحد	٠,٩٠ - ٠,٩٥
- إعادة الاختبار بعد سنة	٠,٨٥
- إعادة الاختبار بعد سنتين ونصف	٠,٨٠
- إعادة الاختبار بعد تسع سنوات	٠,٧٨

٨ - علاقة الموضوعية بالثبات

يرى ستراند وويلسون Strand and wilson ١٩٩٣م أن الموضوعية Objectivity صورة (شكل أو صيغة) Form من صور الثبات، وهى تعرف بالثبات المقدر Rater Objectivity وهو ثبات يحدث عندما يقوم شخصان أو أكثر بتطبيق نفس الاختبار على نفس الأفراد ثم يحصلون على نفس النتائج.

ويمكننا القول بأن موضوعية الاختبار تتأثر تأثرا كبيرا عندما تختلف درجات المحكمين Judges عندما يقومون بتقدير درجة موضوع واحد، حيث تصبح الدرجات التى يتم الحصول عليها من تطبيق الاختبار لا يعتد بها كدرجات تتمتع بالثبات والصدق. فالفروق فى النتائج تحدث عندما يستخدم المحكمون أساليب مختلفة للحكم أو التقويم أو تقدير الدرجة، وأيضا عندما لا يلتزمون بإجراءات وتعليمات تطبيق الاختبار.

مفهوم الموضوعية:

تشير الموضوعية إلى مدى إمكانية الحصول على درجة صحيحة عندما يقوم فردين أو أكثر بتطبيق الاختبار على نفس الأفراد فى مناسبتين (موقفين) مختلفين. والموضوعية تشبه الثبات ولكنها تستلزم قيام اثنين أو أكثر بتطبيق نفس الاختبار وتقدير درجاته.

ويتفق علماء القياس على أن الموضوعية Objectivity عكس الذاتية Subjectivity، فمن المعروف أن المعلم فى المدرسة يعتمد على أسلوب التقدير الذاتى كأسلوب سهل ومعتاد - فى معظم نظم التعليم - لتقدير درجات الاختبارات (الامتحانات) وبخاصة التى تستخدم الورقة والقلم.

وقد لوحظ أنه عندما يقوم اثنين أو أكثر من المعلمين بتصحيح نفس ورقة الإجابة، فإنه يظهر تباين فى تقديرهم للدرجة، حيث يعزى هذا التباين إلى الفروق فيما بين هؤلاء المعلمين بالنسبة لتقديرهم لدرجة الامتحان، لهذا السبب تتجه البحوث العلمية إلى استخدام أساليب الاختبارات الموضوعية بالنسبة لاختبارات الورقة والقلم بدلاً من الأساليب التقديرية.

مما سبق يتضح أن مصطلح الموضوعية يعنى: موضوعية درجة أداة القياس عندما يقوم شخصان (محكمان) أو أكثر بتقدير درجة أى اختبار يطبق على مفحوص واحد فى أوقات مختلفة، حيث تظهر هذه الموضوعية فى مدى اتساق الدرجات التى يتم الحصول عليها، وأنه عندما تختلف الدرجات التى

يضعها المحكمان اختلافات كبيرة، فإن عملية القياس تصبح أقل موضوعية. وكقاعدة عامة: فإن الاختبار الأكثر موضوعية هو الاختبار الأكثر ثباتاً وصدقاً كوسيلة أو أداة للقياس.

وتعد اختبارات الصواب - الخطأ، والاختيار من متعدد ، نماذج للاختبارات الكتابية (الورقة والقلم) التى تتمتع بدرجة عالية من الموضوعية، بينما يتضح أن اختبارات المقال واختبارات الإجابات القصيرة - Short answer نماذج لاختبارات أقل موضوعية لتأثرها بدرجة أكبر بتحيز القائم على تقدير الدرجة (المحكم).

وفى المجال الرياضى يوجد العديد من وسائل القياس التى تتأثر برأى مقدر الدرجة (المحكم) حتى فى حالات استخدام مقاييس التقدير Rating scales لتقويم الأداء فى بعض الأنشطة كالغطس Diving ، والجمباز Gymnastic وبخاصة الجمباز الأرضى، وفى حالات تقويم القوام Posture ، والأداء فى مواقف اللعب الحقيقية أثناء المنافسات الرياضية وغيرها، حيث يصعب الاتفاق فيما بين العديد من مقدرى الدرجة أو القائمين على تنفيذ الاختبار (المحكمين)، وربما يرجع ذلك إلى عدم فهم المحكمين لطرق وأساليب تقدير الدرجة، مما يجعلهم يلجأون إلى التقديرات الذاتية لتقويم الأداء والتى عادة ما تتأثر بميول واتجاهات هؤلاء المحكمين، فكلما كان هناك قصور فى فهم تعليمات الاختبار، وشروط تطبيقه، كلما أدى ذلك إلى انخفاض الثبات نتيجة عدم اتفاق القائمين على التنفيذ (المحكمين) بالنسبة لتقدير الدرجة على الاختبار. ومن ثم فإنه يلزم - للتغلب على هذه المشكلة - الاهتمام بإعداد تعلميات الاختبارات بشكل محدد وواضح (مقنن) بغية الحصول على معاملات موضوعية وثبات وصدق مرتفعة.

تعريف الموضوعية:

تعرف الموضوعية على أنها «درجة الاتفاق فيما بين مقدِّرى الدرجة».

مفهوم الاختبار الموضوعى:

الاختبار الموضوعى objective test هو أى اختبار تمَّ إعداده بطريقة تضمن أقل درجة من التحيز bias أو الأحكام الذاتية لمقدر الدرجة على الاختبار (المحكم). وكلما كانت إجراءات تقدير الدرجة واضحة ومحددة تحديداً دقيقا، زادت موضوعية الدرجة التى يحصل عليها المفحوص.

كيفية حساب الموضوعية:

يتم تقدير الموضوعية - بشكل عام - وفق الأساليب المتبعة فى تقدير الثبات، حيث تستهدف عملية القياس توفير مجموعتين من الدرجات لكل مفحوص، ثم يستخدم أسلوب الارتباط المناسب لتقدير العلاقة بين هاتين المجموعتين من الدرجات.

ويسفر قياس العلاقة بين مجموعتى الدرجات عن معامل ارتباط يعرف باسم معامل الموضوعية objectivity coefficient وهو معامل ليس له رمز رياضى يميزه ولكنه يكون بديلا لمعامل الارتباط الذى استخدم لحسابه.

ولحساب موضوعية الاختبار (الثبات المقدر) يقوم اثنان من المحكمين بتقدير درجة الاختبار، أو الحكم على الأداء لعدد من المفحوصين (الطلاب أو اللاعبين) فى نفس الوقت، بحيث يكون كل محكم بعيداً عن الآخر حتى لا يتأثر برأيه، ثم يلى ذلك حساب معامل الارتباط بين تقديرات المحكمين، حيث يدل معامل الارتباط المحسوب على درجة الاتفاق فيما بين تقديرات المحكمين الاثنين.

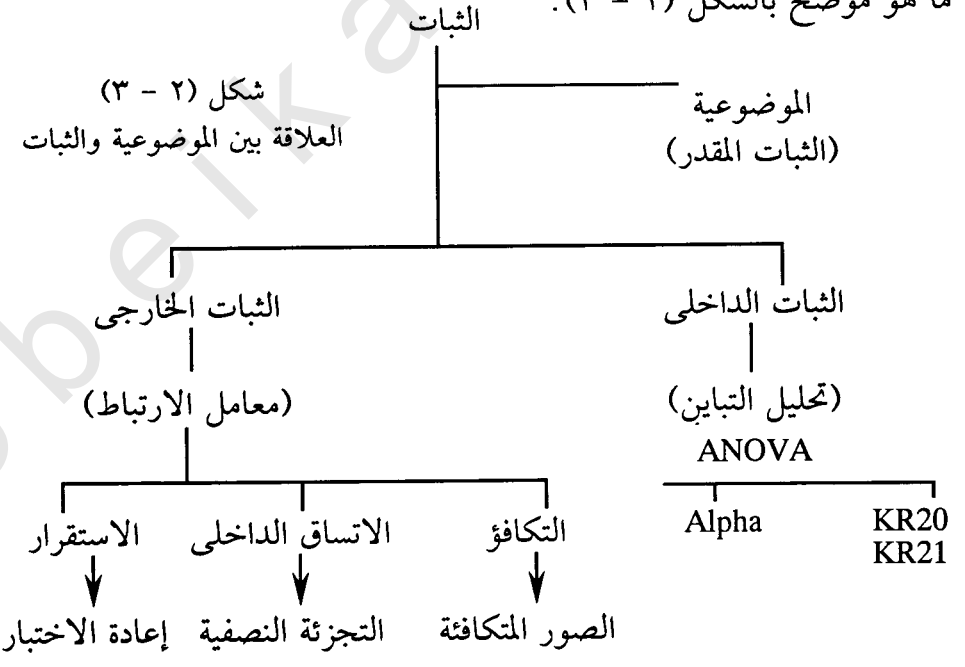
وهناك طريقة أخرى لحساب الموضوعية، إلا أنها طريقة محدودة الاستخدام، وهى طريقة تعتمد على تطبيق الاختبار على نفس مجموعة الأفراد فى موقفين مختلفين، حيث يقوم بتقدير الدرجة فى كل موقف محكم مختلف، يلى ذلك حساب الارتباط بين درجات كل محكم، حيث يلاحظ

أن هذا الأسلوب يقترب من طرق حساب الثبات، ومن ثم نجد أنه أسلوب محدود الاستخدام حيث يغنى عنه تقدير الثبات.

وقد يستخدم تحليل التباين ANOVA لحساب الموضوعية كما هو متبع بالنسبة للثبات، وذلك فى الحالات التى يستخدم فيها تقديرات ثلاثة محكمين أو أكثر. ويتم تقويم موضوعية الاختبارات والمقاييس فى المجال الرياضى وفقا للمعايير التى ذكرها كير كندال وآخرون ١٩٨٧ وهى :

معامل الارتباط (الموضوعية)	مستوى الموضوعية
٠,٩٥ إلى ١,٠٠	ممتاز
٠,٨٥ إلى ٠,٩٤	مرتفع
٠,٧٠ إلى ٠,٨٤	متوسط
٠,٠٠ إلى ٠,٦٩	غير مقبول

ويحدد بومجارتنر وجاكسون ١٩٧٥ م العلاقة بين الموضوعية والثبات وفق ما هو موضح بالشكل (٢ - ٣).



شكل (٢ - ٣)
العلاقة بين الموضوعية والثبات

معجم إنجليزى - عربى للمصطلحات

A

Analysis of variance methocl

طريقة تحليل التباين

Alternate - Forms reliability

معامل ثبات الصور المتبادلة

B

Behavioral disposions

النزوع السلوكى

C

Correlation For attenuation

التصحيح للتنفية

Complex traits

السمات المعقدة

Carry - over effect

انتقال تأثير الأداء

Coefficient of equivalence

معامل التكافؤ

Comparable - forms reliability

معامل ثبات الصور المتساوية

Clinical tests

الاختبارت الاكلينيكية

D

Difficulty value of items

درجة صعوبة الفقرات

Discriminative value

القيمة التميزية للفقرة

E

Equivalent - forms method

طريقة الصور المتكافئة

Error score

درجة الخطأ

Environmental conditions

الاشتراطات البيئية

G

Guessing التخمين

Group Variability تشتت مجموعة الأفراد

H

Human behavior سلوك الإنسان

Heterogeneous عدم التجانس

Hemogeneus التجانس

Homogeneity of items تجانس الفقرات

I

Internal Consistancy method

طريق الاتساق الداخلى

Internal - Consistency reliabitity

ثبات الاتساق الداخلى

ixtrinsic Factors عوامل خارجية

obeikandl.com

Intrinsic factors عوامل داخلية

Item analysis تحليل الفقرة

Intercorrelation الارتباطات البينية

M

Measurement errors أخطاء القياس

Measuring instrument أداة القياس

Motor response الاستجابة الحركية

Momentory fluctuations الأحداث الطارئة

O

Observed score الدرجة الملاحظة (التجريبية)

P

Psychometrics القياس النفسى

Product - moment correlation معامل الارتباط التتابعى

Parallel forms الصور المتكافئة - المتوازية

Parallel - forms reliability معامل ثبات الصور المتوازية

Performance tests الاختبارات العملية (الأدائية)

Paper-and pencil tests اختبارات الورقة والقلم

R

Range of scores مدى الدرجات

Reliability theory نظرية الثبات

Reliabitity Coefficients معاملات الثبات

Reliability of measurements ثبات القياس

Retesting coefficient معامل اعادة الاختبار

Retest reliability ثبات الاختبار عن طريق اعادته

Relative reliability بات النسبى

S

Simple qualities الخصائص البسيطة

Sampling error خطأ المعاينة

Source of error مصادر الخطأ

Split - Half method طريقة التجزئة النصفية

Stability Coefficient معامل الاستقرار

Speed tests اختبارات السرعة

Spelling ability القدرة على التهجئة

T

True variance التباين الحقيقى

total variance التباين العام

Time Sampling المعاينة الزمنية

True score الدرجة الحقيقية

V

Verbal response الاستجابة اللفظية (اللغوية)

Variability الانتشار (التشتت)

obeikandl.com

مراجع الفصل الثانى

١ - سعد عبدالرحمن (١٩٩٨) : القياس النفسى - النظرية والتطبيق. دار الفكر العربى، القاهرة.

٢ - محمد حسن علاوى، ومحمد نصر الدين رضوان (٢٠٠١) : اختبارات الأداء الحركى . دار الفكر العربى، القاهرة.

٣ - محمد حسن علاوى، ومحمد نصر الدين رضوان (٢٠٠٠) : القياس فى التربية الرياضية وعلم النفس الرياضى. دار الفكر العربى، القاهرة

٤ - محمد حسن علاوى، ومحمد نصر الدين رضوان (١٩٨٧) : الاختبارات المهارية والنفسية فى المجال الرياضى . دار الفكر العربى، القاهرة.

٥ - محمد نصر الدين رضوان (١٩٩٧) : المرجع فى القياسات الجسمية، دار الفكر العربى، القاهرة.

6 - Baumgartner, T.A. and Jakson, A.S. (1982) : Measurement for evaluation in physical education. 2 n d ed., Dubuque , Iowa : Wm.C.Brown.

7 - Jackson, A., Jackson, A.S., and Bell,J. (1980) : A Comparison of alpha and the intraclass reliability Coefficients. R.Q., 51,568-571.

8 - Safrit, M.J., and wood, T.M. (1989) : Measurement Concepts in physical education and execise science . champaign, Ill. : Human Kinetics.

9 - Sax, G. (1980) : principles of educational and psychological measurement and evaluation. Belmont, C. A : Wadsworth.

10 - Verducci, F.M. (1980) : Measurement Concepts in physical education. st. Louis : C.V. Mosby.