الفصل الخامس

الثبات

- تعريف الثبات
- أنواع معامل الثبات
- نظرية الثبات
- العوامل المؤثرة في ثبات الاختبار
- أساليب حساب معامل الثبات
- إعادة الاختبار
- الصور المتكافئة
- التجزئة النصفية
- تحليل التباين
- ثبات المصححين

obeikan.com

obeikan.com

الثبات Reliability

نشأ مفهوم الثبات نتيجة لظهور أعداد كبيرة من الأدوات والاختبارات التي تقيس مختلف نواحي السلوك الإنساني وذلك بعد الحرب العالمية الثانية، وبالتالي نشوء شركات القياس التي أخذت تتبنى هذه الاختبارات وتقوم بتسويقها، والتي واجهت مشكلة تتمثل في الأعداد الكبيرة من الاختبارات حيث أصبح لكل مفهوم أو مصطلح نفسي عدد كبير من الاختبارات التي تقيسه، والسؤال الذي يطرح نفسه هو أي من هذه الاختبارات أجود وأفضل من حيث دقته في قياس الظاهرة التي يراد قياسها ؟ وهذا التساؤل جعل شركات القياس تتجه إلى إيجاد المعايير الأساسية التي يحتكم إليها لتحديد جودة الاختبار، وكان من أهم المعايير والتقنيات مفهوم الثبات، حيث بدأ علماء القياس في حسابه بطرق عديدة تقوم الفكرة الأساسية لمعظمها على معامل الارتباط.

ويعد مفهوم الثبات من المفاهيم الجوهرية في القياس، ويمثل مع مفهوم الصدق من الأسس المهمة للاختبارات والمقاييس، إلا أن مفهوم الصدق أشمل وأهم من مفهوم الثبات، إذ يمكن القول إن كل اختبار صادق هو بالضرورة ثابت، لكن ليس كل اختبار ثابت هو صادق بالضرورة. حيث أن الاختبار الصادق الذي يقيس فعلاً ما أعد لقياسه، فإن درجته ستكون معبرة عن الأداء الحقيقي أو القدرة الفعلية للفرد، لذلك مادامت الدرجة على المقياس الصادق تعبر عن هذه الوظيفة بدقة، فإنها تكون ثابتة في الوقت نفسه أي متسقة داخلياً في تعبيرها عن مقدار الوظيفة، ومستقرة عبر الزمن في تعبيرها وتقديرها لهذه الوظيفة إلا أن هذا لا يعني أنه يمكن الاستغناء عن الثبات في إعداد وبناء الاختبارات والمقاييس، وذلك لعدم توفر مقاييس واختبارات تتسم بالصدق التام، بالإضافة إلى أن تقدير الصدق يخضع عادةً لاعتبارات عملية متعددة تختلف من حالة لأخرى ومن مجتمع لأخر مما ينبغي تقدير الثبات أضافه لتقدير الصدق قبل التطبيق النهائي للاختبار أو المقياس على عينة

معينة أو مجتمع معين. ويوفر معامل الثبات Reliabilit coefficient الكثير من المؤشرات الإحصائية للصفة أو الظاهرة المدروسة، والتي من خلالها يمكن الحكم على دقة المقياس الذي استخدم في القياس، بالإضافة إلى أن تقدير ثبات الاختبار يزود الباحث بمعلومات أساسية للحكم على نوعية تكنيك الاختبار ومدى صلاحيته ودقته واتساقه فيما يزودنا به من بيانات عن الصفة أو الظاهرة المدروسة.

وتبرز أهمية الثبات في كوننا لا نملك مقاييس واختبارات سيكولوجية وتربوية مطلقة الأحكام والدقة، كما لا نمتلك التحكم التام والضبط الدقيق لمواقف القياس، مما تتسرب من خلالها بعض المتغيرات الدخيلة المؤثرة في القياس، وبالتالي فإن الدرجة التي نحصل عليها من المقياس أو الاختبار لا تكون معبرة بدقة عن الظاهرة التي نقيسها، سواء أكانت سمة أو قدرة أو غيرهما، إذ تتضمن الدرجة التي نحصل عليها دائماً قدراً من الخطأ سواء أكان خطأً موجباً في شكل زيادة في الدرجة عما يستحقه الفرد نتيجة لقدرته الحقيقية، أم خطأ سالباً في شكل نقصان في الدرجة، مما يكون تقييم أداء الفرد أقل من الواقع الذي يستحقه. وإذا حصلنا على درجات دقيقة وصحيحة بواسطة أي اختبار أو مقياس، بدون أخطاء يمكن أن نرجعها إلى طبيعة ونوعية فقرات الاختبار، أو إلى تخمينات وتحيزات الفرد عند الإجابة على الاختبار، أو إلى الأخطاء التي تحدث بسبب سوء الفهم من قبل الشخص المختبر، أو إلى العوامل الأخرى التي يمكن أن تؤثر على أدائه الحقيقي، فعند ذلك سيمتلك الاختبار ثباتاً تاماً (كامل) مما سيكون معامل ثبات مثل هذا الاختبار يساوي (١).

بيد أن في الواقع، لا يوجد اختبار تحصيلي أو اختبار عقلي أو نفسي أو أي مقياس، يمكن أن يحصل على درجة ثبات تامة (١)، لأنه لا يمكن تجنب جميع هذه الأخطاء، لذلك فإن مهمة الباحث أو المختص هو تقليل الأخطاء المحتومة في القياس إلى أدنى درجة معقولة. وعليه فإن الدرجة التي نحصل عليها من أي اختبار لا تعبر عن الأداء الحقيقي للفرد فقط، بل هي تمثل الأداء الحقيقي للفرد والشوائب الدخيلة في الدرجة سواء كانت موجبة أو سالبة، بمعنى أخر أن درجة الفرد على الاختبار أو المقياس تعبر عن التباين الحقيقي للفرد True variance وتباين الخطأ

Error variance. ولذلك يؤكد (جيلفورد) guilford على ضرورة حساب ثبات الاختبار كي نحدد الدرجة الحقيقية أو التباين الحقيقي للاختبار، لأن معامل الثبات يوضح نسبة التباين الحقيقي في الدرجة المحسوبة على الاختبار وعليه فإن مهمة أساليب حساب الثبات المختلفة هي أن توفر تقديراً جيداً لحجم التباين الحقيقي في الدرجة الكلية المستخرجة مع الإشارة إلى تباين الخطأ.

تعريف الثبات:

يقصد بثبات الاختبار أن تكون أدوات القياس على درجة عالية من الدقة والإتقان والاتساق والاطراد فيما تزودنا به من بيانات عن السلوك المفحوص.ويرى علام انه متى ما كانت أداة القياس خالية من الأخطاء العشوائية، وكانت قادرة على قياس المقدار الحقيقي للسمة أو الخاصية المراد قياسها قياسا متسقا وفي ظروف مختلفة ومتباينة كان المقياس عندئذ مقياساً ثابتاً. ولهذا فإن الثبات هو الاتساق والدقة في القياس. ويمكن أن نوضح مفهوم الثبات من خلال مقابلته مع مفهوم الصدق حيث أن مفهوم الثبات يعني الاتساق في مجموعة درجات الاختبار التي فعلاً قاست ما يجب قياسه. أما مفهوم الصدق فيعني الدقة في درجات الاختبار التي يفترض أن تقيس ما يجب قياسه. أما التعريف الإجرائي لمعامل ثبات مجموعة من درجات مجموعة من المتحنين فهو معامل ارتباط بين تلك المجموعة من الدرجات ومجموعة أخرى من الدرجات على اختبار مكافئ، تم الحصول عليها بشكل مستقل من أفراد نفس المجموعة وهذا التعريف الإجرائي يتطلب قياسين أو أكثر لنفس الصفة أو الظاهرة لكل عضو من أعضاء المجموعة بواسطة اختبارات متكافئة ويستخدم مفهوم الثبات بمعناه العام ليدل على مدى اعتماد الفروق الفردية في درجات الاختبار على أخطاء الصدفة المتضمنة في القياس من ناحية، وعلى الاختلافات الحقيقية في الصفة أو الخاصية موضع الدراسة من ناحية أخرى. والثبات بهذا المعنى يشير إلى نسبة التباين الحقيقي في الدرجة التي يحصل عليها الفرد من اختبار ما، ولذلك فإن معامل الثبات هو هذه النسبة من تباين الدرجة على المقياس التي تشير إلى الأداء الفعلي للمفحوص، مما يتضمن هذا المعنى تصنيف الدرجة التي

يحصل عليها الفرد إلى مكونين رئيسين هما التباين الحقيقي لأداء الفرد على المقياس وتباين الخطأ أو الدرجة الزائفة نتيجة لشوائب القياس. وهناك بعض المؤشرات والملاحظات التي ينبغي ذكرها كي تكون الصورة واضحة عن معنى وأهمية تقدير الثبات أو الاختبار، ومنها:

١ـ أن الثبات ضروري، لكنه ليس هو الحالة الكافية لتحديد نوعية ودقة الاختبار، حيث قد نحصل على درجات ثبات عالية، لكننا قد نقيس بدقة شيئاً ما لا علاقة له بالموضوع أو الظاهرة المراد قياسها. أما أذا حصلنا على درجات غير ثابتة (Unreliable) فإن المقياس لا ينجح في قياس أي شيء تماماً من الظاهرة أو الصفة المدروسة. مما ينبغي أن نتأكد أولاً من موضوعية وصدق المقياس في قياس السمة أو الظاهرة المراد قياسها ومن ثم التعرف على درجة ثباته.

٢ ـ إن الاختبار أو المقياس نفسه بوصفه مجموعة من الفقرات أو الأسئلة لا ثبات له، بينما الثبات بالأحرى إلى الاختبار عندما يطبق على مجموعة أو عينة من الأفراد ولذلك فإن الثبات هو ثبات الإجابة على الاختبار، أي ثبات أداء الفرد على الاختبار، ولهذا فإن ما يحسب ثباته هو عينة استجابات مجموعة من الأفراد.

٣ ـ إذا كانت الفروق بين درجات الفرد نفسه ذات صلة قليلة بالفروق بين درجات الأفراد المختلفين فإن الاختبار يتجه إلى إعطاء ثبات عالي، في حين إذا كان الفرق بين درجات الفرد نفسه ذات صلة كبيرة بالفروق بين الأفراد فإن الدرجات تظهر ثباتاً واطئاً.

٤ ـ بما أن الثبات يعني ثبات عينة الاستجابات من مجموعة الأفراد، وبما أنه لكل مجموعة من الأفراد خصائصها، فمن الضروري أن نحصل على أكثر من معامل ثبات للاختبار الواحد لدى أكثر من عينة.

٥ ـ إن حصولنا على معامل ثبات لأداء عينة من الراشدين على اختبار معين، لا يعني أن ثبات أداء المراهقين أو الأطفال سيكون في نفس المستوى من الثبات، كما لا يجوز مطلقاً أن نقبل ثبات أداء عينة من الذكور على أنه مطابق لثبات أداء عينة من الإناث على نفس الاختبار. كما تلعب الفروق الحضارية والتعليمية وغيرها دوراً هاماً في التأثير على ثبات المقاييس أو الاختبارات. لذلك ينبغي أن نقدر معامل ثبات

أي اختبار قبل تطبيقه على عينة الظاهرة المدروسة مهما كان ثابتاً على عينات أخرى أو في بيئة أخرى أو في وقت سابق.

٦ ـ تتعدد أساليب حساب معامل الثبات ويختص كل أسلوب أو طريقة منها لتقدير نوعية معينة من تباين الخطأ ، وبالتالي نوعية معامل الثبات المراد تقديره. وعلى الباحث أن يحدد نوع معامل الثبات قبل استخدام الأسلوب الذي يقيس ذلك النوع من الثبات والجدول التالي يوضح أنواع معامل الثبات وتباين الخطأ وطريقة تقدير كل نوع منها.

	نوع معامل الثبات	طريقة التقدير	تباين الخطأ
١ ١	معامل الاستقرار	أعادة نفس الاختبار على العينة نفسها بعد فترة زمنية	التغيرات المؤقتة
٢ ٢	معامل التكافؤ	تطبيق الاختبار بصورتيه المتكافئتين في نفس الوقت	عينة محتوى الاختبار
٣ ٣	معامل الاستقرار والتكافؤ	تطبيق الاختبار ثم تطبيق الصورة المكافئة له بعد فترة مناسبة	التغيرات المؤقتة وعينة المحتوى
٤ ٤	معامل الاتساق الداخلي	التجزئة النصفية للاختبار، ثم يصحح بمعادلة (سبيرمان – بروان) أو غيرها.	عدم تجانس نصفي الاختبار
٥ ٥	معامل الاتساق داخل الأسئلة	تحليل الفقرات باستخدام أحد معادلتي (كيودر – ريتشارد سون) أو غيرهما.	التجانس الكلي لفقرات الاختبار

نظرية الثبات:

تتولى نظرية الثبات مهمة تحليل طبيعة كل من التباين الحقيقي وتباين الخطأ والعلاقة بينهما لتوفر المنطق الأساسي الذي تقوم عليه الطرق المختلفة لحساب ثبات الاختبار، حيث أن:

الدرجة الكلية المستخلصة من المقياس = التباين الحقيقي لأداء الفرد على المقياس + تباين الخطأ نتيجة لشوائب المقياس ..... معادلة (١)

أما إذا استخدمنا مقياس محكم تماماً في اختبار فرد ما تحت ظروف مثالية للقياس فسنجد أن التباين الحقيقي هو عبارة عن الدرجة التي يمكن أن يحصل

عليها هذا الفرد، غير أن هذا المستوى من القياس وهذه الظروف لا توجد في الواقع العملي مما لا يمكن أن تكون الدرجة الكلية ممثلة للتباين الحقيقي لدرجة الفرد. والطريقة الأخرى التي يمكن أن تلجأ أليها لتحديد الدرجة الصحيحة للفرد (أي التباين الحقيقي) هي أن نختبر نفس الفرد عدداً كبيراً من المرات بنفس المقياس وتحت نفس الظروف أو ظروف مشابهة ثم نحسب متوسط الدرجة التي يحصل عليها الفرد في هذه الاختبارات المتعددة وبهذا تكون الدرجة الحقيقية أو التباين الحقيقي عبارة عن متوسط درجات القياسات المستقلة بواسطة نفس الاختبار.

والمنطق الذي يؤيد اختبار متوسط الدرجة في هذه الاختبارات المختلفة هو التباين الحقيقي هو أن تباين الخطأ أو الجزء الزائف من درجة الفرد على المقياس قد يكون سالباً أو موجباً، ووفقاً لخصائص المنحني الأعتدالي فإن متوسط هذه الدرجات أو المقادير الخاطئة في عدد كبير من مرات الاختبار سيكون صفراً. وبذلك يعبر متوسط أداء الفرد في هذا العدد الكبير من الاختبارات عن درجته الحقيقية لأن الدرجات الزائدة الموجبة والدرجات الناقصة السلبية ستتعادل وتصبح صفراً وتبقى فقط الدرجة المعبرة عن أدائه الفعلي أي المتوسط. إن هذا المنطق سليم نظرياً من ناحيتين:

الأولى:

إن المجموع الجبري لتباين الخطأ سيكون صفراً في عدد كبير للغاية من مرات الاختبار للفرد الواحد وإن كان من المحتمل أن لا يكون هذا المجموع صفراً في عدد قليل من مرات القياس، غير أن هذا لا يقلل من منطق توزيع تباين الخطأ على جانبي المنحنى الأعتدالي

الثانية: إن الدرجة الحقيقية للفرد وفقاً لهذا المنطق هي الأكثر استقراراً بينما التغيير يحدث في الجزء الزائف من الدرجة الكلية، وهذا يعني بتعبير أخر نتيجتين محدودتين هما:

١ ــ إننا لا نتوقع ارتباطاً بين الدرجة الحقيقية والدرجة الزائفة أو التباين الحقيقي وتباين الخطأ فأحدهما متغير والأخر ثابت.

٢ ـ إننا لا نتوقع أيضاً وجود ارتباط بين الدرجات الخاطئة في صورة ما من اختبار والدرجات الخاطئة في صورة متكافئة لنفس الاختبار، لأن الدرجات الزائفة عشوائية الاتجاه وغير منتظمة ولا يحكمها منطق القدرة الحقيقية التي يقيسها الاختبار.

وهذا المنطق يؤدي إلى صياغة عدد من المعادلات الجديدة التي تترتب على معادلة (١) التي مر ذكرها، والتي توضح هذه المعادلات طبيعة تباين الخطأ والتباين الحقيقي:

معادلة (٢): متوسط تباين الخطأ = صفر

لأن تباين الخطأ الموجب والسالب يتوزع إعتدالياً.

معادلة (٣):

الارتباط بين التباين الحقيقي وتباين الخطأ = صفر

على افتراض استقرار الأداء الحقيقي وتذبذب تباين الخطأ إيجاباً وسلباً بين المرة والأخرى من مرات القياس.

معادلة (٤): الارتباط بين الدرجات الخاطئة لصورتي نفس الاختبار = صفر

على افتراض عشوائية اتجاه تباين الخطأ في كل مرة من مرات الاختبار وبين جزئي الاختبار الواحد.. وهي النتيجة التي تعتمد عليها كل أساليب التصنيف في حساب الثبات. لقد اعتمدنا في ترتيب المعادلات السابقة من خلال موقف افترضنا فيه أننا نقوم باختبار الشخص الواحد عشرات المرات اختباراً مستقلاً وبنفس المقياس، غير أن هذا الإجراء غير ممكن عملياً لأسباب كثيرة، ولذلك نعود مرة أخرى من المعادلة (١) ومن المعادلات الأخرى المتسقة معها لنضع عدداً من الصيغ الجبرية التي تتعلق بالمكونات الصحيحة والخاطئة للمتوسط والتباين الخاص بالدرجة التي نحصل عليها من المقياس. وبما أن مجموع متوسطات أي عدد من المكونات الموحدة تساوي متوسط هذه المجاميع فيمكننا أن نضع المعادلة الآتية:

معادلة (٥): متوسط الدرجات الكلية = متوسط التباين الحقيقي + متوسط تباين الخطأ .

ولما كان متوسط تباين الخطأ = صفراً (معادلة ٢) ولذلك فإن متوسط الدرجات الكلية سيساوي متوسط التباين الحقيقي. ولما كان منطقياً أن متوسط مجموعة التباينات يساوي مجموع متوسطات هذه المجموعة من التباينات فنستطيع أن نضع معادلة للتباين كما وضعت للمتوسط وهي

معادلة (٦): التباين الكلي = التباين الحقيقي + تباين الخطأ

وبهذه المعادلة نكون قد وصلنا إلى عدد من المفاهيم التالية، التي تمكننا من الانتقال إلى الوسائل الإجرائية لاستخلاص الثبات أو التباين الحقيقي للاختبار، والمفاهيم هي:

١ ـ إن متوسط تباين الخطأ يساوي صفر.

٢ ـ لا يوجد ارتباط بين المقادير الخاطئة.

وبذلك يكون الثبات هو ((النسبة من التباين الصحيح في الدرجات المستخلصة من الاختبار)) ويمكن أن نعبر عن هذه النسبة بالمعادلة آلاتية : **معادلة.. (٧)**

$$\text{ثبات الاختبار} = \frac{\text{التباين الحقيقي}}{\text{التباين الكلي}}$$

وعلينا الآن أن نتعرف على حجم التباين الحقيقي لكي نحسب ثبات الدرجة على الاختبار، وهذا ممكن إذ اتضح من المعادلة (٦) أن هناك ارتباطاً بين التباين الكلي ومكونات الاختبار (التباين الحقيقي وتباين الخطأ) وهذا الارتباط يتيح لنا الحصول على صيغة جديدة للمعادلة (٦) بالمعادلة الآتية:

معادلة (٨) التباين الحقيقي = التباين الكلي – تباين الخطأ

وبالتعويض عن هذه الفروق في المعادلة (٨) مع المعادلة (٧) نحصل على **معادلة.. (٩)**

$$\text{معامل الثبات} = 1 - \frac{\text{تباين الخطأ}}{\text{التباين الكلي}}$$

ومن مميزات هذه المعادلة (٩) هي أن البيانات التجريبية توفر لنا المعلومات اللازمة لحلها، حيث نستطيع أن نحصل على حجم التباين الكلي وتباين الخطأ وبالتالي نحسب معامل الثبات.

حساب التباين الحقيقي: يمكن من خلال صيغة المعادلة رقم (٧) أن نحصل عملياً على تقدير للتباين الحقيقي لدرجات الاختبار، فإذا كان التباين الكلي ومعامل الثبات معروفين فتحل المعادلة (٧)

$$\text{معامل الثبات} = \frac{\text{التباين الحقيقي}}{\text{التباين الكلي}}$$

تصبح لدينا معادلة رقم (١٠)

معادلة رقم (١٠) التباين الحقيقي = معامل الثبات × التباين الكلي.

وبإيجاد الجذر التربيعي للتباينات في المعادلة السابقة يصبح لدينا تقدير للانحراف المعياري للدرجات الصحيحة.

$$\text{الانحراف الحقيقي} = \sqrt{\text{الانحراف المعياري} \times \text{معامل الثبات}}$$

أما مصادر الخطأ التي تؤدي إلى أخطاء القياس تتمثل في الآتي:

١ـ مستوى الفرد في سمة عامة معينة يؤثر على أدائه لعدد من المقاييس.

٢ ـ مدى فهم تعليمات المقياس وطريقة أدائه.

٣ ـ دافعية الفرد للإجابة عن المقياس.

٤ ـ التزييف والكذب والخداع من قبل الفرد.

٥ ـ الظروف البيئية مثل التهوية ودرجة الحرارة والضوضاء.

٦ ـ تعود الفرد على أداء نوع معين من الاختبارات والمقاييس.

٧ ـ الاضطرابات العقلية البسيطة مثل اضطراب الذاكرة والانتباه.

٨ ـ دور عامل الصدفة والتخمين في اختيار الإجابات الصحيحة.

obeikan.com

وإذا ما استطعنا التعرف على بعض هذه المصادر فإنه يمكن علاجها وبالتالي يزيد ثبات المقياس، حيث نقلل من تأثيرها ليبقى التباين الموجود بين درجات الأفراد يعكس الفروق الفردية الحقيقية في الصفة أو السمة المراد قياسها لدى هؤلاء الأفراد. والخلاصة من كل ما سبق يمكن إيجازها في أن المقياس يزيد ثباته كلما تحققت النقاط التالية:

١ـ الحصول على نفس النتائج تقريباً عند كل مرة يتم فيها إعادة التطبيق.

٢ـ ان يكون التباين الحقيقي اكبر ما يمكن بالنسبة للتباين العام او أن تباين الخطأ اقل ما يمكن.

٣ـ ان يكون هناك علاقة بينية وارتباطية بين وحدات وأجزاء المقياس. لان وجود مثل هذه العلاقة يعطي دلالة على تناسق البناء الداخلي للمقياس وبالتالي فإن معامل الثبات يعتمد على هذه العلاقة البينية بين كل وحدة وأخرى وقيمة الارتباط بين كل وحدة وبين المقياس ككل. حيث إن تماسك المقياس وتناسق بنائه يدل على ثباته بل انه يمكن حساب معامل الثبات من خلال هذه العلاقة بين وحدات المقياس.

العوامل المؤثرة في ثبات الاختبار:

هناك عوامل عديدة يمكن أن تؤثر على ثبات الاختبار. وبالتالي ينبغي من الباحث أن يتعرف عليها ويعمل على التحكم بها قدر الإمكان وتخفيض تدخلها في الثبات، ومن هذه العوامل:

أولاً: طول الاختبار (عدد الفقرات أو الأسئلة): يؤثر طول الاختبار أو عدد فقراته في درجة ثباته، إذ كلما تزايد عدد الفقرات كلما أرتفع ثبات الاختبار، وذلك لأن العدد الأكبر من الفقرات يؤدي إلى الحصول على عينة أكبر من السلوك، وبالتالي يكون من المتوقع أن يمثل بهذه العينة بشكل مستقر العدد الأكبر من مكونات السلوك أو السمة أو القدرة المقاسة والقابلة للظهور في مرتي التطبيق أو في نصفي أو جزئي الاختبار. في حين أن الفقرات القليلة في المقياس تقلل من احتمال استقرار مكونات السلوك أو الصفة المراد قياسها، وبالتالي انخفاض معامل الثبات. وكلما كان الاختبار طويلاً كلما زادت فرصة ان الخطأ يلغي بعضه بعضاً، وبالتالي زادت

قيمة التباين مما يؤدي إلى زيادة معامل الثبات. إن تباين الخطأ ثابت والتغير في التباين الكلي يعود إلى تباين الدرجة الحقيقية، وبالتالي عندما يزيد تباين الدرجة الحقيقية فإن التباين الكلي يزداد، حيث ان زيادة عدد فقرات الاختبار تعطي فرصة اكبر لرصد تباين الدرجة الحقيقية وبالتالي زيادة التباين الكلي، وهو ما تؤكده الحقيقة التالية: (إذا زاد طول الاختبار (ن) مرة فإن التباين الحقيقي لدرجته يزيد (ن)^{٢} مرة ويزيد تباين الخطأ (ن) مرة ، وبالتالي زيادة التباين الكلي يؤدي الى معامل ثبات اكبر). وللوصول إلى معامل الثبات المرغوب للاختبار ينبغي إضافة عدد أخر من الفقرات إلى الاختبار، ويمكن استخدام معادلة ((سبيرمان - براون)) في تحديد عدد الفقرات التي ينبغي أضافتها للوصول إلى معامل الثبات المرغوب. كما يمكن استخدام هذه المعادلة في تقليل عدد الفقرات بمعامل ثبات معين، حيث أن معادلة ((سبيرمان - براون)) تساعد على التنبؤ بالزيادة المنتظرة في معامل الثبات إذا أضفنا إلى الاختبار عدداً من الفقرات المشابهة للفقرات الموجودة فيه، وبعبارة أخرى تمكن هذه المعادلة من تحديد العلاقة بين الزيادة في طول الاختبار والزيادة في معامل الاختبار مثلاً، إذا زاد عدد الفقرات من ٢٥ إلى ٧٥ فقرة، فعندها يصبح عدد فقرات الاختبار (٣) مرات عددها قبل الزيادة، فتصبح ن= ٣ أما إذا اختصرنا الأسئلة أو الفقرات من (١٠٠) فقرة إلى (٥٠) فقرة، فتصبح ن= ½، ومن ثم نستخدم معادلة (سبيرمان ـ براون) التالية:

$$ر_{ن} = \frac{ن \times ر_{س}}{(ن - ١) ر_{س} + ن}$$

حيث أن: ر$_{ن}$: معامل الثبات التقديري بعد زيادة فقرات الاختبار **ن**: عدد مرات إطالة الاختبار أو اختصاره، **ر$_{س}$**: معامل الثبات الذي حصلنا عليه قبل الزيادة أو النقصان. فإذا كان معامل الثبات قبل الزيادة (٠,٥٠) فإن تطبيق المعادلة يكون بالشكل الأتي:

$$ر_{ن} = \frac{(٣) (٠,٥٠)}{١ + (٣-١) (٠,٥٠)}$$

وفي حالة نقصان الفقرات كما في المثال السابق وكان لدينا معامل الثبات)٠,٨٠(فإن تطبيق المعادلة يكون بالشكل الأتي:

$$ر_{ن} = \frac{(٠,٨٠)\ (٠,٥٠)}{١ + (٠,٨٠)\ (١ - ٠,٥٠)}$$

ويمكن اختصار معادلة (سبيرمان - براون) السابقة فيما إذا أردنا أن نتنبأ بمعامل ثبات اختبار يبلغ طوله ضعف طوله الأول كما هو الحال في حساب الثبات بالتجزئة النصفية ولذلك تصبح معادلة (سبيرمان - براون) كما يأتي:

$$ر_{٢} = \frac{٢\ ر_{٢}}{١ + ر_{س}}$$

حيث أن: $ر_{٢}$: ثبات الاختبار المصحح لنصفي الاختبار. ٢: عدد مرات الإطالة للاختبار - كما في التجزئة النصفية إلى مرتين. $ر_{س}$: معامل الثبات قبل التصحيح بطريقة معامل الارتباط. لكن ينبغي أن نشير هنا أن تطويل الاختبار ليس هو فقط الطريقة الوحيدة المفضلة لتحسين ثباته إذ أن المعرفة الصحيحة والتكنيك الدقيق للاختبار سيعطي معامل ثبات عالي.

ثانياً: مدى تقارب صعوبة الفقرات ودقتها:

إن وجود فقرات شديدة الصعوبة في الاختبار لا يستطيع جميع الأفراد أو أغلبهم الإجابة عنها لا تضيف ميزة للاختبار ولا يؤثر حذفها في الدرجة التي يحصل عليها أي فرد من أفراد العينة ما دامت لا تميز بين فرد وأخر، وبالمثل فإن وجود فقرات شديدة السهولة التي يستطيع الإجابة عنها كل أفراد العينة فإنها لا تميز بين فرد وأخر، لذلك ينبغي حذف أو تعديل الفقرات الشديدة الصعوبة أو السهولة إذا أردنا أن نرفع ثبات الاختبار لأنها لا تضيف شيئاً لقيمة درجة الثبات، لأن معامل الثبات مبني في الأساس على التباين الحقيقي بين درجات أفراد العينة. وعليه فإن أفضل فقرات في الاختبار هي التي يكون احتمال الإجابة عليها ٥٠ ٪ من الأفراد، كما أن

الفقرات التي قد تحتمل تفسيرات عديدة ومتشعبة فإنها ستؤدي إلى انخفاض تباين الأسئلة وبالتالي انخفاض التباين الكلي للاختبار.

القيمة المقبولة لمعامل الثبات: معامل الثبات عبارة عن نسبة تباين، ومعامل الثبات المقبولة تختلف من اختبار لآخر وذلك حسب الغرض من الاختبار ودقة القرار المترتب عليه وعلى ذلك:

١ـ الاختبارات المقننة تتطلب معاملات ثبات لا تقل عن (٠,٨٠) وذلك عندما تكون القرارات على مستوى الأفراد، أما على مستوى الجماعات فيقبل معامل ثبات (٠,٦٥).

٢ ـ معاملات ثبات الاختبارات التحصيلية المقننة يجب ألا تقل معاملات ثباتها عن (٠,٨٥) بينما يمكن أن تقل عن ذلك في اختبارات الشخصية.

٣ ـ الاختبارات التحصيلية والتي لا يراعى في تطبيقها الخطوات التي يتم مراعاتها في الاختبارات المقننة معظمها تتراوح معاملات ثباتها بين (٠,٢٠ ـ ٠,٤٠) ونادرا ما تصل إلى (٠,٦٠)

ثالثاً: تعلق الإجابة على فقرة بالإجابة على فقرات أخرى: عند استخدام اختبار يتضمن عدد من الفقرات ترتبط الإجابة عليها بعضها ببعضها الأخر يؤدي إلى انخفاض معامل ثبات الاختبار، لأن هذه الخاصية ستؤدي إلى خفض عدد الفقرات، مما ينبغي أن تكون كل فقرة مستقلة في نوعية الإجابة عليها حتى وإن كانت تقيس سمة واحدة أو قدرة واحدة.

رابعاً: موضوعية التصحيح: تعتبر موضوعية التصحيح من العوامل المؤثرة في ثبات الاختبار، ولا سيما في الاختبارات التي تعتمد على تقدير المصحح كاختبارات المقال التحصيلية واختبارات الإبداع أو الاختبارات الأسقاطية، حيث أن تباين التصحيح يؤدي إلى زيادة تباين الخطأ وبالتالي إلى نقصان معامل الثبات. مما ينبغي اللجوء إلى تقدير معامل ثبات المصححين، كما ينبغي أن يتضمن الاختبار وصفاً دقيقاً لمحكات الإجابة والتصحيح وطريقة وضع الدرجة عليها.

خامساً: أثر تخمين المجيب: في الاختبارات التي تعتمد على تقدير الفقرة أو السؤال بعدد من البدائل، يلجأ بعض المفحوصين عادةً في حالة عدم تأكدهم من الإجابة الصحيحة إلى التخمين واختيار أي أجابه منها، وعادةً ما يكون التخمين في

مثل ذلك عند إعادة الاختبار أوفي جزء أخر من الاختبار، مما يؤدي الى خفض ثبات الاختبار ويمكن معالجة أثر التخمين أو الكذب لا سيما في اختبارات الشخصية في إحدى الطرق التالية:

١ ـ من الممكن أضافه بعض الفقرات الكاشفة للكذب بين فقرات الاختبار أو المقياس والتي من خلالها نتمكن من معرفة دقة الإجابات وعدم الكذب في الإجابة وبالتالي يتمكن الباحث من إهمال اختبار الفرد الذي تكون درجة الكذب لديه عالية في المقياس أو الاختبار. ويستخدم هذا الأسلوب عادة في اختبارات أو مقاييس الشخصية واختبارات القدرات المميزة.

٢ ـ وضع بعض الضوابط التي تجبر المفحوص على التفكير على الإجابة قبل أن يؤشر عليها كأن تكون الفقرة مكونة من عبارتين لهما نفس المقبولية الاجتماعية لكنهما تختلفان في قياس الظاهرة، ويحدث هذا في اختبارات القدرات المميزة أيضاً.

٣ ـ أما في الاختبارات التحصيلية فيمكن استخدام المعادلة الآتية بمعالجة أثر التخمين:

$$\text{د ح} = \frac{\text{د ص} - \text{د خ}}{\text{ن} - ١}$$

حيث أن: د ح = الدرجة الحقيقة، د ص=عدد الفقرات الصحيحة، د خ = عدد الفقرات الخاطئة ،ن = عدد بدائل الإجابة على الفقرة

سادساً: زمن الاختبار: يؤثر الزمن المحدد للإجابة على الاختبار على ثباته، فيزداد الثبات تبعاً لزيادة الزمن حتى يصل إلى الحد المناسب للاختبار - أي الزمن الذي تتطلبه الإجابة على الاختبار - فيصل الثبات إلى أعلى درجة ممكنة من جراء تأثير هذا العامل. أما إذا كان الاختبار يتطلب وقتاً كبيراً لإنجازه، فإن ذلك سيؤدي إلى خفض معامل الثبات، لأن ذلك قد يؤدي إلى الإجهاد والتعب وضعف الدافعية على الإجابة وبالتالي قد تتعرض الفقرات الأخيرة من الاختبار إلى أخطاء في الإجابة ومن ثم تؤدي إلى زيادة تباين الخطأ. وعليه ينبغي تحديد الزمن المطلوب للاختبار ووفقاً للجهد المطلوب للإجابة عليه ومدى قدرة المفحوص على الاستمرار في

الإجابة دون تعب، وتحديد مدة الاختبار المطلوبة يمكن الوصول إليها عن طريق تجريبه على عينات استطلاعية.

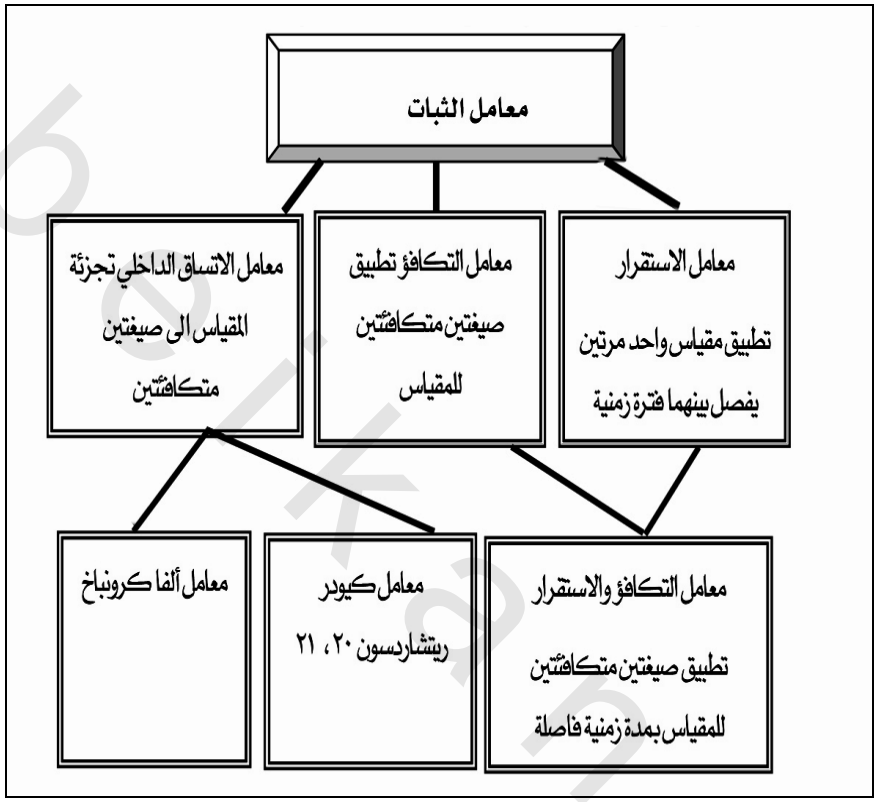
سابعاً: تجانس العينة:يؤدي التجانس الشديد في عينة الثبات من حيث الظاهرة المدروسة إلى انخفاض واضح في معامل الثبات، لأن التباين داخل هذه العينة المتجانسة يكون منخفضاً بقدر لا يسمح بتقدير التباين الحقيقي - أي ثباته - في حين كلما كبر حجم العينة وتزايد عدم تجانسها كلما كانت أقرب إلى التوزيع الإعتدالي وبالتالي سيصبح تباين الخطأ صغيراً مما تؤدي هذه النتيجة إلى ارتفاع ثبات الاختبار. ويمكن تفسير ذلك أيضاً إلى أن تجانس أفراد العينة سوف يؤدي إلى انخفاض تباين كل سؤال، حيث أن تباين السؤال يعتمد على نسبة الأفراد الذين أجابوا إجابة صحيحة ونسبة الأفراد الذين أجابوا إجابة خاطئة، وبالتالي انخفاض تباين الأسئلة يؤدي إلى انخفاض التباين الكلي للاختبار، ومن المعلوم أن تباين الخطأ ثابت وبالتالي فإن قسمة تباين الخطأ على التباين الكلي سوف يكون كبيرا وعند طرحه من الواحد الصحيح للحصول على معامل الثبات فإن معامل الثبات يكون منخفضاً.

ثامناً: بعض العوامل الأخرى التي تؤثر بمقادير مختلفة في ثبات الاختبار: كضبط موقف التطبيق، ودافعيه المفحوص والمؤثرات الفيزيقية والمشتتات المتعددة في موقف الاختبار، والحالة الصحية والانفعالية للمفحوص، مما ينبغي أن يقوم الباحث بضبط دقيق لكل هذه العوامل حتى يتمكن من الوصول الى معاملات ثبات مرتفعة.

أساليب حساب معامل الثبات:

تتعدد وتتنوع أساليب حساب الثبات في الاختبارات والمقاييس، ويختص كل أسلوب منها بتقدير نوعية محددة من (تباين الخطأ) وهو التباين الذي يؤثر في ثبات القياس الذي نحصل عليه كلما استخدمنا مقاييس مختلفة. وتتميز بعض الاختبارات والمقاييس بإمكان حساب ثباتها بأسلوب دون آخر وعدم صلاحية بعض الأساليب لها، إلا أن هذا لا يعني أن المقياس أو الاختبار الواحد لا يصلح له إلا أسلوب واحد من أساليب حساب الثبات حيث أن في بعض الحالات يمكن أن تستخدم أكثر من أسلوب واحد للاختبار الواحد بهدف التعرف على مصادر تباين

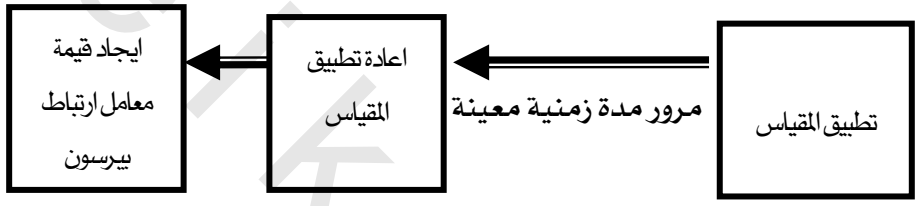
الخطأ الذي يؤثر في استقرار أو أتساق الدرجة التي نحصل عليها من الاختبار. وفيما يلي استعراض موجز لأساليب حساب الثبات ونوعية الثبات التي يقيسها كل أسلوب وطريقة حسابه ويمكن تلخيص هذه الطرق بالمخطط الآتي:



أولاً: إعادة الاختبار: Test – retest

يسمى معامل الثبات الناتج من هذه الطريقة بمعامل الاستقرار coefficient of stability، أي استقرار نتائج الاختبار خلال الفترة بين التطبيق الأول والثاني للاختبار. حيث أن هذا النوع في حقيقته هو معامل الارتباط بين مجموعتي درجات الاختبار في التطبيقين، وذلك بإعطاء نفس مجموعة الأفراد الاختبار نفسه مرتين، ومصدر الخطأ في الدرجة التي يحصل عليها الفرد في ضوء معامل الثبات هو خطأ التغيرات عبر الفترة الزمنية. ولا يستخدم هذا الأسلوب في استخراج الاتساق الداخلي، أذ قد يكون معامل الارتباط عالياً بينما يكون معامل الثبات الداخلي

منخفضاً. ويفضل عادةً عند حساب الثبات بهذا الأسلوب آلا يكتفي الباحث بحسابه على مدى فترة زمنية واحدة بل على أكثر من فترة، إذ من الممكن أن يعاد التطبيق بعد أسبوعين ومن ثم أعادته بعد فترة شهر ثم أعادته مرة ثالثة بعد فترة أخرى. ويستفاد من هذه النتائج في إمكانية الباحث بالتنبؤ بنتائج اختباره عبر الزمن، كما يمكن للباحث أن يستخدم أكثر من عينة متجانسة تماماً ويعيد عليها الاختبار بعد فترة مختلفة كي يتخلص تقريباً من احتمال تأثر أداء العينة بإعادة الاختبار سواء كان هذا التأثر إيجابياً أو سلبياً نتيجة للمران أو التدريب والنمو أو نتيجة للملل أو عدم الاهتمام.



ويواجه هذا الأسلوب (إعادة الاختبار) في إيجاد معامل الثبات بعض الانتقادات والعيوب، لعل من أهمهما ما يأتي:

١ـ تمثل الفقرات التي استخدمت في الاختبارين عينة واحدة فقط. والتي تكون عادةً من مجتمع واحد من الفقرات المتيسرة للاختبار، مما لا تكون درجات إعاده الاختبار دليلاً واضحاً عن كمية تغير مقدار الدرجات فيما إذا استعملنا عينة مختلفة من الفقرات أو الأسئلة.

٢ـ إن إجابات المفحوص في الاختبار الثاني لا تكون مستقلة عن إجاباته على الاختبار الأول، إذ قد تتأثر استجاباتهم في المرة الثانية في تذكرهم لاستجاباتهم في التطبيق الأول، أو من خلال المناقشات التي تجرى بين المفحوصين بعد الانتهاء من التطبيق الأول، أو قد تتأثر بواسطة أنواع أخرى من الجهود التي يبذلها بعض المفحوصين عادةً لمعرفة الإجابة الصحيحة بعد الانتهاء من التطبيق الأول مما تساعدهم هذه الجهود في معالجة الصعوبات التي واجهوها وبالتالي ستتحسن إجاباتهم في التطبيق الثاني.

٣ـ قد تتداخل أخطاء القياس مع التغيرات الحقيقية في قابليات وقدرات المفحوصين عند إعادة الاختبار نتيجة للتعلم أو النمو.

٤ـ في أحيان كثيرة بعض المفحوصين لا سيما من الطلبة لا يروق لهم استخدام هذا الأسلوب باعتباره مضيعة لوقت التعلم وإرباك للنظام المدرسي لا سيما في التطبيق الثاني، كما أن المعلمين لا يحبذون ضياع وقت الطلبة في مثل هذه الأمور، مما يؤدي ذلك إلى ضعف رغبة قسم من الطلبة في الإجابة على الاختبار الثاني أكثر من الأول وبذلك يقل ثبات الاختبار.

٥ـ التذبذب العشوائي في أداء الفرد الواحد في الموقعين، لأن أداء الفرد لا يتم من خلال قوالب صارمة في كل الأوقات، وبالإضافة إلى التغيرات التي قد تحدث في حالته الصحية أو النفسية أو تغيرات الجو وما إلى ذلك من مؤثرات على الفرد أثناء الاختبار مما تؤثر بشكل ما على معامل الثبات. ويمكن للباحث أن يقلل من مساوئ إعادة الاختبار عن طريق إعداد صورتين متكافئتين من اختبار معين، تعطى الصورة الأولى في وقت معين وتعطي الثانية بعد ذلك فترة، ثم يحسب معامل الارتباط بين درجات الصورتين لإيجاد معامل الثبات وبذلك يتأكد الباحث بهذه الطريقة من استقرار الاختبار وتكافئه.

شروط استخدام أسلوب (أعاده الاختبار) في أيجاد الثبات: أن استخدام هذا الأسلوب في إيجاد ثبات الاختبار أو المقياس محكوم ببعض الشروط والقواعد التي ينبغي على الباحث مراعاتها والتأكد منها قبل استخدام أسلوب (إعادة الاختبار) في إيجاد معامل الثبات، لعل من أهمها ما يأتي:

أولاً: أن تتسم الظاهرة المدروسة بنوع من الاستقرار، وليست متذبذبة.

ثانياً: أن يبتعد الباحث عن استخدام هذا الأسلوب في الاختبارات التي قد يؤدي التطبيق الأول لها إلى نوع من التعلم أو التدريب على الاستجابات الصحيحة، أو التي تخضع إلى التذكر، لذلك فإن هذا الأسلوب لا يصلح للاختبارات والمقاييس التي تهدف إلى قياس التذكر. أو التي تؤدي الإجابة عليها إلى نوع من التدريب والتعلم، كاختبارات التحصيل وبعض الاختبارات التي ترتبط بالعمليات العقلية، في حين يمكن استخدامه في اختبارات الشخصية بعد مراعاة الفاصل الزمني المفضل بين

التطبيقين، وتصلح للاختبارات ذات الزمن المحدد إلى حد كبير، والاختبارات ذات الزمن غير المحدد لكونه يقيس قوة الاستجابة أكثر من سرعتها.

ثالثاً: أن يعمل الباحث على توفير نفس الظروف للمفحوصين في الموقف الأختباري، الأول والثاني، لأن الاستجابات قد تتأثر بالظروف المحيطة بالمفحوصين أكثر مما تتعلق بالمقياس أو الاختبار.

رابعاً: على الباحث مراعاة الفاصل الزمني بين التطبيق الأول والثاني للاختبار، والتي لا يمكن تحديد فترة ملائمة لجميع المقاييس والاختبارات، إذ أن طول هذه الفترة يعتمد عموماً على طبيعة الظاهرة المدروسة وطبيعة عينة البحث حيث أن درجة الثبات أو معامل الاستقرار يتأثر بطول الفترة الزمنية بين التطبيقين للاختبار، فإذا كانت قصيرة جداً فهناك احتمال كبير أن يتذكر المجيب إجاباته في التطبيق الأول، في حين إذا كانت الفترة طويلة جداً فقد يحصل تغير حقيقي في الظاهرة أو الصفة المدروسة مما يرتفع الثبات في الحالة الأولى ويقل في الحالة الثانية.

لذلك يفضل أن تكون الفترة الزمنية قصيرة في الحالات التالية:

أ- إذا كانت الظاهرة تتأثر بالنمو الجسمي والعقلي لأفراد العينة مثلاً وفي الممارسة العملية يتطلب أعادة الاختبار في أيجاد الثبات لها فترة قصيرة بعد التطبيق الأول للاختبار.

ب ـ عند اختبار الأطفال الصغار وبسبب سرعة معدل النمو لديهم، لأن التغيرات التي قد تحدث في الأداء النسبي بعد فترة زمنية طويلة يمكن اعتبارها ضرباً من النمو والتقدم وليس مجرد تغيير عشوائي ناتج عن الظروف التي قد تطرأ على الموقف الاختياري التي لا يمكن التحكم فيها.

جـ ـ في الاختبارات التي لا تتأثر الاستجابات في التطبيق الثاني بتذكر استجابات التطبيق الأول لها.

في حين يفضل أن تكون الفترة الزمنية طويلة في الحالات التالية:

أ- إذا كانت الظاهرة المدروسة لا تتأثر بالنمو والتغيير السريع.

ب- في حالة كون أفراد العينة كبار السن ولديهم نوعاً من استقرار النمو.

جـ ـ إذا كانت الاستجابات تتأثر بالتذكر من التطبيق الأول للاختبار.

وعموماً فإن تحديد الفترة الفاصلة بين التطبيقين للاختبار يتوقف على دراسة طبيعة الظاهرة وعينة التطبيق، وقد تمتد هذه الفترة ما بين نصف ساعة إلى ستة أشهر وأحياناً أكثر، في حين يرى البعض أن لا تقل هذه الفترة عن أسبوع ولا تزيد عن ستة أشهر، وكما سبق أن بينا أن تحديد الفترة من قبل الباحث لا يتم اعتباطاً بل حسب طبيعة البحث ونوعية العينة وخصائص الظاهرة المدروسة.

خامساً: أما كيفية حساب معامل الثبات بهذا الأسلوب فيتم بالخطوات الآتية:

أ- يقدم الاختبار أو المقياس إلى أفراد العينة ثم يصحح وتدون نتيجة أو درجة كل فرد.

ب- يعاد تطبيق الاختبار على أفراد العينة بعد الفاصل الزمني الملائم وضمن ظروف مشابهة في التطبيق الأول، ثم يصحح حسب قواعد تصحيح التطبيق الأول له ثم تدون النتائج أو الدرجات لكل فرد أمام درجته في التطبيق الأول.

جـ ـ يتم حساب معامل الارتباط بين مجموعتي الدرجات - درجات التطبيق الأول ودرجات التطبيق الثاني - ومعامل الارتباط هذا يمثل معامل الثبات، وعلى الباحث أن يراعي طبيعة البيانات عند استخدام معامل الارتباط، فإذا كانت متصلة أي غير منفصلة فيمكن استخدام (معامل ارتباط ـ بيرسون) أما إذا كانت البيانات من النوع المنفصل فيفضل استخدام (معامل ارتباط - سبيرمان للرتب) إلا أنه في حالة وجود درجات أو قيم مماثلة كثيرة، أي أن نفس الدرجة تتكرر لدى عدد من الأفراد فإنه لا يفضل استخدام (معامل ارتباط سبيرمان) بل يفضل استخدام (معامل ارتباط كاندل).

وتعد هذه الطريقة مناسبة لاختبارات الاستعداد وذلك لان نتائج اختبار الاستعداد تستخدم في التنبؤ بعيد المدى عن مستقبل الطالب التربوي أو المهني، وبالتالي فإنه من الضروري أن نتعرف على مدى استقرار درجات الاستعداد، كما تعتبر مناسبة لحساب معامل ثبات الاختبارات الموقوتة ذات الزمن المحدد والتي تعتمد إلى حد كبير على السرعة، كما أنها تلائم مقاييس الاتجاهات والميول اعتمادا على أن الاتجاهات والميول بطيئة التغيير، ونجد إن هذه الطريقة لا تصلح للاختبارات

التي تهدف إلى قياس التذكر أو ترتبط به ارتباطاً مباشراً، كما أنها لا تناسب اختبارات التحصيل لأن التحصيل سريع التغيير.

ثانياً:- أسلوب الصور المتكافئة Equivalent forms

إذا أعددت صورتان متماثلتان أو أكثر لاختبار واحد بنفس الطريقة فمن المحتمل أن درجات هذه الصور البديلة سوف تظهر متكافئة، وحينها يتم إعطاء كل فرد من أفراد العينة صورتي الاختبار، فإن الارتباط بين الصورتين سيكون معاملا للثبات. ويسمى معامل الثبات بهذا الأسلوب بمعامل التكافؤ coefficient of equivalence حيث يدل على مدى الارتباط بين درجات اختبار واختبار أخر متكافئ معه يعطى في الوقت نفسه أو من وقت متقارب جداً. وتباين الخطأ في هذا الأسلوب يدل على أخطاء عينة المحتوى - الأسئلة أو الفقرات - ويعد هذا الأسلوب من أفضل الأساليب في حساب الثبات في الاختبارات التحصيلية في حين لا يصلح للقياس في الاختبارات أو المقاييس الموقوتة.

ويمتاز أسلوب الصور المتكافئة ببعض المميزات التي تجعله أحياناً مفضلاً في حساب معامل الثبات ومن هذه المميزات ما يأتي:

١- يساعد هذا الأسلوب في خفض احتمال تأثير التدريب أو التعلم والتذكر على النتائج كما في أسلوب (إعادة الاختبار).

٢- إذا تم استخدام الصورتين المتكافئتين في وقتين مختلفين، فإننا نضمن تلافي بعض العيوب (إعادة الاختبار) كما يمكن ان تقيس به معامل الاستقرار والتكافؤ معاً.

٣- معامل الثبات باستخدام الصور المتكافئة، فيه تشابه كبير مع معامل الاتساق الداخلي الذي يتم حسابه بأسلوب التجزئة النصفية، لا سيما في تطبيق الصورتين في الوقت نفسه أو متقاربة جداً.

١- تعد الصور المتكافئة من الأساليب المناسبة لتقدير مدى تمثيل عينة الفقرات للمجال الذي يقيسه الاختبار، فكلما كان هذا التمثيل جيداً، فإننا نتوقع

أن ينعكس على الصورة المكافئة، وهنا يظهر أيضاً التجانس في أداء الفرد على الصورتين في شكل معامل الثبات.

إلا أن هذه الميزات لا تمنع من وجود بعض الصعوبات في هذا الأسلوب، منها صعوبة إعداد صورتين متكافئتين تماماً لها نفس المحتوى ونفس الهدف فيما تقيسه من الظاهرة المدروسة وما تحتاج من جهد ووقت وتكلفة عند إعدادهما. ويمكن تلخيص طريقة حساب معامل التكافؤ بالمخطط الآتي:

قواعد وشروط أعداد الصور المتكافئة:

١ ــ أن يبذل الباحث جهداً واضحاً للوصول إلى فقرات متكافئة في صور الاختبار سواء أكان ذلك في صياغتها أو مستوى صعوبتها أو تمثيلها للسمة المقاسة مع استخدام نفس التعليمات، لكنها ذات استقلالية أي أنها غير متداخلة أومتشابهة في بعض الفقرات.

٢- ويشير (ثورندايك) إلى أن الصور المتكافئة ينبغي أن يكون لها نفس التباين الحقيقي، كما ينبغي أن لا يحدث تداخلاً في تباين الخطأ فيها، أذ كلما كان الفرق بين التباين الحقيقي كبيراً كلما صغر معامل الثبات، وكلما تداخل تباين الخطأ كلما زاد معامل الثبات، لأن معامل الارتباط الذي يحسب بينهما يعتمد على تشابه التباين الحقيقي وتباين الخطأ التجريبي ويرى (جلكسون) أن الصور المتكافئة لها نفس المتوسط ونفس التباين، كما يفضل استخدام ثلاث صور للاختبار ومن ثم حساب معاملات الارتباط بينها.

٣- عند استخدام معامل الارتباط بين درجات العينة لكلا الصورتين، على الباحث أن يراعي طبيعة الدرجات لاستخدام نوع معامل الارتباط المناسب - كما سبق توضيحه في أسلوب أعاده الاختبار.

ثالثاً: أسلوب التجزئة النصفية Spilt – halves:

يعتمد هذا الأسلوب أساساً على تقسيم فقرات المقياس أو الاختبار الى قسمين متكافئين، ثم حساب معامل الارتباط بين درجات هذين القسمين ولذلك فإن هذا الأسلوب يصلح في الاختبارات التي تكون فيها فقرات المقياس أو الاختبار متجانسة أي أنها تقيس خاصية أو سمة واحدة وعليه فإنه يقيس التجانس بين الفقرات، أي الاتساق الداخلي، مما يسمى معامل الثبات بهذا الأسلوب (معامل الاتساق الداخلي) coefficient of consistency وأن تباين الخطأ الذي يقيسه أسلوب التجزئة النصفية هو عدم تجانس نصفي الاختبار. ومن الواضح أن حساب الثبات بالتجزئة النصفية لا يقيس التجانس الكلي للمقياس لأنه يقسم الفقرات إلى قسمين، لذلك فهو معامل ثبات لنصف الاختبار أو المقياس، ولتلافي ذلك ينبغي تصحيح معامل الثبات المستخرج بطريقة حساب الارتباط بين درجاتها باستخدام بعض الطرق الإحصائية لإيجاد معامل الثبات لكل الاختبار، ومنها معادلة (سبيرمان - براون) Spearmen - Brown formula كما هناك بعض الطرق التي يحسب بها كل معامل الثبات والتي سيمر ذكرها فيما بعد. كما أن هذا الأسلوب يجنب الباحث بعض الصعوبات التي تواجهه في أيجاد معامل الثبات بأسلوب (إعادة الاختبار) وبأسلوب (الصور المتكافئة) ويمكن اعتباره من أكثر الأساليب عملية وسهولة. كما ينبغي أن نؤكد أن تقسيم الاختبار إلى جزأين لا يتم قبل تطبيقه ولا يؤثر عليه، بل يقدم على شكل متكامل، ومن ثم يجري تقسيمه بعد ذلك، وهناك عدة طرق لتقسيم الاختبار أو المقياس إلى قسمين، من أهمها الآتي:

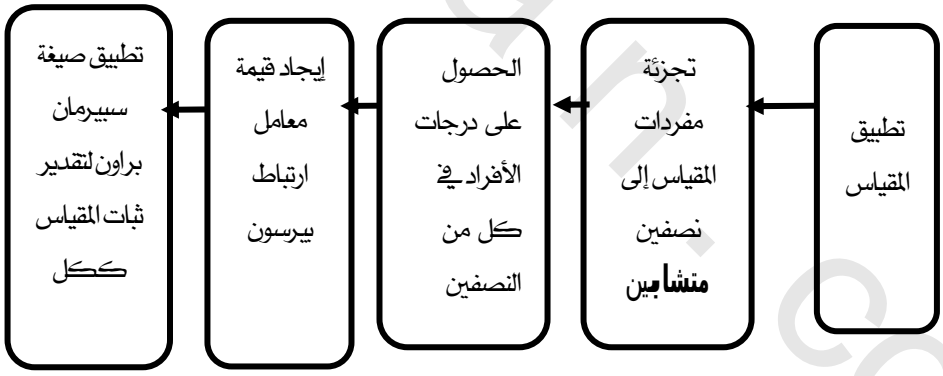
١ ـ القسمة النصفية، وذلك بتقسيم فقرات الاختبار إلى نصفين متساويين يمثل النصف الأول المجموعة الأولى من الفقرات ويمثل النصف الثاني المجموعة الثانية من الفقرات، بيد أنه لا يفضل استخدام هذه الطريقة لاعتبارات عديدة منها أن المجيب قد لا يكون بنفس الدافعية أو القدرة عند الإجابة على النصف الثاني من الاختبار.

٢ ـ الفقرات الفردية والفقرات الزوجية، وهذه من الطرق الشائعة في التجزئة النصفية.

٣ ـ جزءا الاختبار: رغم أن طريقة التصنيف الى فردي وزوجي من أفضل الطرق التي قد تؤدي إلى نصفين متعادلين في خصائصها ومتساويين في تعرضها لظروف الأداء، إلا أن بعض المقاييس والاختبارات لا تصلح هذه الطريقة في تقدير ثباتها لا سيما أذا كانت تقدم كل مجموعة من الفقرات في وحدة زمنية معينة أو يكون المطلوب من المبحوثين تقديم عدد كبير من الاستجابات لفترة واحدة في فترة زمنية معينة ثم تقدم الفقرة الأخرى وهكذا، ولذلك يفضل تقسيم الفقرات إلى (٤) أجزاء مستقلة كل جزأين في قسم واحد، وقد يتكون

القسم الأول من الجزأين (١، ٣) والقسم الثاني من الجزأين (٢، ٤). وهناك من يقسم الفقرات إلى قسمين بعد تحديد مستوى صعوبة كل فقرة ومن ثم توزيعها بالتساوي على قسمي الاختبار، كما يمكن توزيعها بناء أعلى مدى في صلاحيتها للتمييز بين الأفراد، وهذا الإجراء يتطلب عادةً تطبيق الاختبار على عينة استطلاعية بهدف تحديد مستوى صعوبة وتمييز كل فقرة.

ويمكن توضيح ذلك بالمخطط الآتي:



شروط استخدام أسلوب التجزئة النصفية:

١ ـ كما يشير (جلكسن) إنه من الخطأ اعتبار التجزئة النصفية على أنها صورتان متكافئتان، ولكي تكون كذلك ينبغي أن تكون متعادلتين من حيث المتوسط ودرجة الصعوبة والتشتت ومعاملات الارتباط بين الوحدات.

٢ ـ لا تصلح إلا في اختبارات القوة لا السرعة، كي تتساوى الفقرات المستخدمة في حساب معامل الارتباط، ويكون هناك قدر كاف من الفقرات لحساب معامل الارتباط.

٣ ـ تستخدم في المقاييس والاختبارات التي تكون فيها الفقرات متجانسة أي أنها تقيس خاصية أو سمة واحدة.

طرق حساب الثبات لأسلوب التجزئة النصفية: معادلة (سبيرمان – براون) (Spearman – Brown) بما أن حساب الثبات بالتجزئة النصفية، هو عبارة عن ثبات نصف الاختبار وليس كله، لذلك ينبغي تصحيح معامل الارتباط الذي يمثل معامل الثبات في التجزئة النصفية بمعادلة (سبيرمان - براون) حيث تساعد هذه المعادلة على حساب معامل الثبات إذا أردنا أن نضيف فقرات إلى الاختبار بعد حسابه بطريقة معامل الارتباط، أي أنها تساعد في تحديد العلاقة بين الزيادة في معامل الثبات وعدد الفقرات التي ينبغي أن تضاف إلى الاختبار. بيد أن هناك بعض الانتقادات الموجهه لهذه المعادلة في تصحيح معامل الثبات الذي تم إيجاده بحساب معامل الارتباط، منها:

١- افتراضها تشتت واحد لنصف الاختبار، ونادراً ما يكون الانحراف المعياري متساوياً في النصفين.

٢- تحتاج إلى عمليات حسابية كثيرة، لأنها تحتاج إلى حساب معامل الارتباط بين جزئي الاختبار.

٣- مبالغتها في رفع قيمة معامل الثبات النصفي.

٤- لا تصلح مع الاختبارات الموقوتة.

٢. معادلة (رولون) Rulon:

تعتبر معادلة (رولون) طريقة مختصرة لحساب معامل الثبات الكلي للاختبار ولا تحتاج إلى أيجاد معامل الارتباط بين جزئي الاختبار كما في معادلة (سبيرمان – براون) وهي:

$$ر = ١ - \frac{ع_{د}^{٢}}{ع_{س}^{٢}}$$

حيث أن: ر : معامل الثبات، $ع_{د}^{٢}$: تباين الفرق بين درجات نصفي الاختبار،

$ع_{س}^{٢}$: التباين الكلي لمجموع درجات نصفي الاختبار

وتتجه معادلة (رولون) إلى تقدير الثبات بالاعتماد على نسبة التباين الحقيقية للأداء فقط مؤكدة أن الفرق بين التباين المنطقي الذي يفترض أن يعبر عنه الاختبار هو (١) وبين تباين الخطأ هو معامل الثبات الفعلي للاختبار. وعلى هذا فالمتوقع في حالة الثبات المطلق للاختبار آلا يوجد فرق بين أداء الفرد على نصفي الاختبار ولهذا تستخدم معادلة (رولون) هذا الفرق بين الدرجات على نصفي الاختبار لتقدير ثباته. كما أن هذه المعادلة تحسب ثبات الاختبار كله دون حاجة لإعادة تصحيح الطول، ويرجع ذلك إلى حقيقة أن الفرق بين درجتي الفرد على نصفي الاختبار عبارة عن مجموع تباين الخطأ في النصفين معاً وليس نصفا واحداً فقط. ولهذا فالفروق بين النصفين خاصةً بالاختبار كله، أي أن التباين الذي تعبر عنه هذه الفروق هو الجزء من التباين في الدرجة الكلية الراجع للخطأ في كل الاختبار وليس تباين خطأ نصف الاختبار.

معادلة موزير Mosier: تتميز معادلة موزير بأنها مختصرة، إذ أنها توفر القدر الأكبر من الجهد المبذول في العمليات الحسابية، لكنها تتطلب التصحيح بمعادلة (سبيرمان - براون) ونص المعادلة كالأتي:

$$ر_{ف ز} = \frac{(ر_{ف ك})(ع_{ك}) - ع_{ف}}{ع^{٢}_{ك} + ع^{٢}_{ف} - (٢)(ر_{ف ك})(ع_{ف})(ع_{ك})}$$

حيث أن: $ر_{ف ز}$: الارتباط بين النصف الفردي والزوجي.

$ر_{ف ك}$: الارتباط بين النصف الفردي مثلاً والاختبار كله.

$ع_{ف}$: الانحراف المعياري لدرجات النصف الفردي.

$ع_{ك}$: الانحراف المعياري لدرجات الاختبار كاملاً.

٣- معادلة جتمان:

معادلة (جتمان) تعتمد على نفس المنطق الذي اعتمدت عليه معادلة (رولون) والفارق بينهما لا يخرج عن كونه فارقاً في أسلوب الحساب الذي يجعل من معادلة (جتمان) أكثر سهولة، إذ يستخدم فيها التباين الخاص بكلا النصفين حيث تقوم بجمع هذا التباين وقسمته على التباين الكلي للاختبار، كما لا تحتاج معادلة (جتمان) لاستخدام معادلة أخرى لتصحيح الطول، ونص المعادلة كالأتي:

$$ر_{ك} = ١٢ - \left(\frac{ع_{١}^{٢} + ع_{٢}^{٢}}{ع_{ك}^{٢}} \right)$$

حيث أن: $ر_{ك}$: ثبات الاختبار. $ع_{١}^{٢}$: تباين الجزء الأول من الاختبار.

$ع_{٢}^{٢}$: تباين الجزء الثاني من الاختبار. $ع_{ك}^{٢}$: تباين الاختبار كله.

٣- معادلة هورست Horst:

عندما تمنع بعض الاعتبارات العملية أحياناً التصنيف المتوازن للاختبار كأن تقوم بقسمة الاختبار إلى جزئين غير متعادلين، مما لا يصح هنا استخدام أي من المعادلات السابقة ويمكن فقط استخدام معادلة (هورست)، ونص المعادلة الآتية:

$$ر_{ك} = ر + \frac{\sqrt{[ر^{٢} + ٤ (ك) (ص) (١ - ر^{٢}) - ر]}}{٢ (ك) (ص) (١ - ر^{٢})}$$

حيث أن: $ر_{ك}$: ثبات الاختبار كاملاً ر: الارتباط بين جزئي الاختبار.

ص : النسبة الصغرى من الاختبار الممثلة للجزء الأصغر.

ك : النسبة الكبرى من الاختبار الممثلة للجزء الأكبر.

وتعد معادلة (هورست) معادلة تصحيح طول مثلها في ذلك مثل معادلة سبيرمان -براون، ما دامت تبدأ من حساب الارتباط بين جزأي الاختبار غير المتساويين وتنتهي إلى أعادة تقدير هذا الارتباط مدخلة في اعتبارها الطول المختلف لكل جزء.

معادلة جلكسون Guliksen:

معادلة التنبؤ (لسبيرمان - براون) تتأثر بالزمن المحدد للاختبار وبذلك لا تصلح مع حساب ثبات الاختبارات الموقوتة التي قد تحول دون إكمال الاختبار في الوقت المحدد للإجابة، لذلك اقترح (جلكسون) المعادلة الآتية لحساب ثبات الاختبارات الموقوتة:

$$ر''_{١١} = ر_{١١} - \frac{م_{ت}}{ع^{٢}_{خـ}}$$

حيث أن: ر" ١١ : معامل ثبات الاختبار الموقوت أو معامل الثبات بعد تصحيح أثر السرعة.

ر ١١: معامل الثبات المحسوب بطريقة (سبيرمان - براون).

م ت : متوسط الأسئلة المتروكة في آخر الاختبار (وذلك بجمع الأسئلة المتروكة عند جميع الأفراد وتقسيم المجموع على عدد الأفراد).

ع٢ خ: تباين الخطأ (ويحسب برصد عدد الاستجابات الخاطئة عند كل فرد مع إضافة عدد الأسئلة المحذوفة ثم يحسب تباينها بالنسبة لكل الأفراد).

٣- **معادلة (فلانجان)** Flanagan:

$$ر_{11} = \frac{٤ ع_{1} \times ع_{2} \times ر_{21}}{ع_{1}^{2} + ع_{2}^{2} + ٢ ع_{1} ع_{2} \times ر_{21}}$$

حيث أن: ر ١١: معامل الثبات للاختبار كله.

ع١ : الانحراف المعياري للجزء الأول من الاختبار.

ع٢ : الانحراف المعياري للجزء الثاني من الاختبار.

ر ٢١: معامل الارتباط بين جزئي الاختبار.

ع١٢ التباين للجزء الأول من الاختبار.

ع٢٢: التباين للجزء الثاني من الاختبار.

تستخدم معادلة (فلانجان) في أيجاد معامل الثبات لجزئي الاختبار (للاختبار كله) بعد أيجاد معامل الارتباط بين درجات جزئي الاختبار، وبذلك تستخدم هذه المعادلة للتصحيح كما في معادلة (سبيرمان - براون) لكنها تختلف عنها في العمليات الحسابية.

رابعاً: تحليل التباين Variance Analysis:

إن طريقة التجزئة النصفية تصلح لحساب معامل الاتساق الداخلي Coefficient consistency (تجانس نصفي الاختبار) لكنها لا تقيس تجانس الأسئلة ذاتها في الاختبار. إلا أن بعض المختصين يرون أن تجانس الأسئلة لا يدخل في مجال موضوع (ثبات الاختبارات) وإنما هي خاصية من خصائص الاختبارات الجيدة، ومع ذلك لا يمكن استبعاده من مناقشة الثبات، حيث أنه يتناول استقرار

استجابات المفحوص على أسئلة أو فقرات الاختبار واحداً بعد الأخر، وعادةً ما يسمى هذا النوع من الثبات بـ (معامل الاتساق داخل أسئلة أو فقرات الاختبار) Interitem consistency ويختلف (معامل الاتساق داخل الأسئلة) عن (معامل الاتساق الداخلي) حيث أن معامل الاتساق داخل الأسئلة يشير إلى التجانس الكلي للاختبار بالمعنى المباشر، في حين يكون معامل الاتساق الداخلي للاختبار أقرب إلى معنى التكافؤ كما يتمثل في الصور المتكافئة للاختبار. فقد يكون لدينا اختباران لهما معاملان للثبات عاليان تم حسابهما بطريقة الصور المتكافئة أو بالتجزئة النصفية، لكنهما قد يختلفان في درجة تجانس الأسئلة التي تتكون منها صورتا الاختبار أو نصفاه. وعموماً فإن الفرق بين معامل الاتساق الداخلي - الذي يحسب بالتجزئة لنصفية - وبين معامل الاتساق داخل الأسئلة، يمكن اختباره دليلاً على مدى التجانس أو التباين في أسئلة الاختبار.

إن أشهر الطرق لحساب معامل اتساق الأسئلة هي طريقة (كيودر-ريتشاردسون) kuder - Richardson حيث قاما في سنة ١٩٣٧ بإيجاد معادلة لحساب الثبات هي:

معادلة ($KR_{٢٠}$) وبالشكل الأتي:

$$ر = \frac{ن}{ن - ١}\left(١ - \frac{مج\ ص\ خ}{ع^{٢}}\right)$$

حيث أن: ر: معامل الثبات. ن : عدد فقرات الاختبار.

مج: المجموع بالنسبة لجميع الفقرات.

ص: نسبة الذين أجابوا إجابة صحيحة على الفقرة.

خ: نسبة الذين أجابوا إجابة خاطئة على الفقرة.

$ع^{٢}$: تباين درجات الاختبار.

وهناك معادلة أخرى لـ (كيودر - ريتشاردسون) وتسمى KR٢١ وهي:

$$ر = \frac{ن}{ن - ١}\left(١ - \frac{م\ (ن - م)}{ن\ ع^{٢}}\right)$$

حيث أن: م: الوسط الحسابي لكل الاختبار.

إن المعادلة KR٢٠ قابلة للتطبيق فقط في الاختبارات التي تكون درجة الإجابة على الفقرة أما صحيحة فتأخذ درجة واحدة أو خاطئة فتأخذ صفراً. أما المعادلة KR٢١ فإنها تعطي دائماً تقديراً منخفضاً لمعامل الثبات من أي أسلوب أخر من أساليب التجزئة النصفية، بل وتعطي أكثر انخفاضاً من المعادلة KR٢٠ كما ينبغي مراعاة أن تكون الفقرات متوسطة الصعوبة، أي عدم وجود فقرات صعبة جداً أو فقرات سهلة جداً وإلا أدى ذلك إلى انخفاض معامل الثبات. إن التجانس الداخلي بين الفقرات يمكن ان تقيسها طريقة (كيودر - ريتشاردسون) وتقوم على تقديرها، أما في حالة وجود مقاييس فرعية فإن هذه الطريقة لا تقيس تجانس جميع الفقرات بل تقدير ثبات المقاييس الفرعية المتجانسة.

١- **تعديل (تيكر) لمعادلة (كيودر - ريتشاردسون):**

يهدف تعديل تيكر Tuker الى تبسيط العمليات الحسابية لمعادلة (كيودر-ريتشارسون) KR٢٠ كما يهدف إلى تجنب عدم الدقة في معادلة (كيودر-ريتشاردسون) KR٢١ التي تؤدي إلى ثبات منخفض، ووفقاً لتعديل (تيكر) تصبح المعادلة KR٢٠ كالأتي:

$$ر = \left(\frac{ن}{ن - ١}\right)\left(\frac{ع_{ك}^{٢} - ن\ ص"\ خ" + ن\ ع_{ص}^{٢}}{ع_{ك}^{٢}}\right)$$

حيث أن:

$ع_{ك}^{٢}$: تباين الدرجة الكلية على الاختبار. ص" : متوسط نسبة الإجابات الصحيحة على الفقرة. خ " : متوسط نسبة الإجابات الخاطئة على الفقرة.

ن : عدد فقرات الاختبار.

$$ع_{ص}^{٢} = \frac{مج\ ص^{٢}}{ن} - ص^{٢}$$

حيث أن ص^{2} = مربع نسبة الإجابة الصحيحة على الاختبار، وكما يلاحظ من هذه المعادلة فإن تعديل (تيكر) ينصب على إضافة متوسط نسبة تباين الإجابات الصحيحة على الاختبار في معادلة (كيودر - ريتشاردسون) kr٢١ مما يؤدي إلى تعديل نتيجتها لتطابق معادلة kr٢٠، ولذلك يفضل استخدام تعديل (تيكر) بدلاً من معادلتي (كيودر - ريتشاردسون) (KR٢٠، KR٢١).

٣٠تعديل (دريسيل) لمعادلتي (كيودر - ريتشاردسون)

بينما نجد أن معادلتي (كيودر - ريتشاردسون)، ومعادلة (تيكر) تقبل التطبيق في الحالات المعتادة بالنسبة للاختبارات المتجانسة والتي تصمم لقياس وظيفة واحدة فقط، لذلك فإنها غير مناسبة للاختبارات التي تتضمن نظاماً للتصحيح يقوم على أوزان متدرجة من الفقرات. حيث يمكن استخدام معادلتي (كيودر - ريتشاردسون) في حالة الإجابة على الفقرة إما مقبولة أو مرفوضة، صواب أو خطأ، نعم أو لا، أما إذا كانت الإجابة متدرجة على أوزان معينة فإن المعادلة المناسبة هنا لحساب الثبات بنفس منطق معادلتي (كيودر - ريتشاردسون) أو ثبات الاتساق الداخلي وتجانس الفقرات هي معادلة (دريسيل) والتي نصها:

$$ر_{ن} = \left(\frac{ن}{ن - ١}\right)\left(\frac{ع^{٢}_{ك} - مجـ و^{٢}_{ق} ص_{ق} خ_{ق}}{ع^{٢}_{ك}}\right)$$

حيث أن: و$_{ق}$: الوزن الخاص بالاستجابات الصحيحة للفقرة (ق). ص$_{ق}$: نسبة الإجابات الصحيحة للفقرة (ق). خ$_{ق}$: نسبة الإجابات الخاطئة للفقرة (ق).

وعموماً فإن معادلتي (كيودر ـ ريتشاردسون) أو تعديلاتها لا تصلح لاختبارات السرعة الموقوتة وبذلك تكون مقتصرة على اختبارات القوة فقط.

٨ ـ معادلة (كرونباخ العامة) للثبات:

يقدم (كرونباخ) Cronbach معادلة عامة تعد المعادلات السابقة جزءاً منها، أو حالة خاصة منها، وهي تنطلق من المنطق العام لثبات الاختبار ويطلق على معادلة (كرونباخ) أسم معامل (ألفا) Alpha، والتي صيغتها:

$$\text{معامل الفا} = \left(\frac{\text{ن}}{\text{ن} - ١}\right)\left(١ - \frac{\text{م ع}_{\text{ق}}}{\text{ع}_{\text{ك}}}\right)$$

حيث أن: ع ق: تباين الجزء ق من الاختبار - مهما كان طول هذا الجزء.

ع ك: التباين الكلي للاختبار. ن : عدد أجزاء الاختبار.

وتنطبق هذه الصيغة لحساب الثبات سواء أكانت أجزاء الاختبار عبارة عن نصفين أو تتعدد إلى أن تكون فقراته جميعها كما في معادلة (كيودر - ريتشاردسون). ويفضل استخدام هذا المعامل عندما يكون الهدف تقدير معامل ثبات المقاييس الجوانب الوجدانية والشخصية نظرا لأنها تشتمل على مقاييس متدرجة لا يوجد بها إجابة صحيحة وأخرى خاطئة. ويجب ألا يفوتنا أن نذكر بأن هناك طرقاً أخرى لقياس معامل الثبات في حالة ما إذا كانت الإجابات متعددة وليست ثنائية ولعل أهمها وأشهرها الطرق التي تعتمد على تحليل التباين والتحليل العاملي.

ثبات المصححين Scorer Reliability:

من المصادر الرئيسة لتباين الخطأ في درجة الاختبار أو المقياس الذي يعتمد على تقييم المصحح وليس على مفتاح للتصحيح أو مجرد عند الاستجابات، ما يطلق عليه عدم ثبات التقييم أو التصحيح، وهذا ما يحدث عادةً في الاختبارات الاسقاطية التي تعتمد أساساً على تأويل استجابة المفحوص للمنبه وفقاً لمحاكاة معينة، لكن انطباقه لتأويل على شروط المحاكات كثيراً ما يرجع إلى ذاتية المصحح رغم التزامه بشروط المحك، مما تظهر الفروق بوضوح عندما يستخدم مصححان أو أكثر نفس المحك لتصحيح نفس العينة من الاختبارات، كما تظهر نفس الحالة أيضاً في الاختبارات الخاصة بقياس القدرات الإبداعية، حيث أن المصحح هو الذي يحدد مدى جودة أو أصالة الاستجابة في اختبارات الأصالة رغم توفر معايير التصحيح. كما أن الاختبارات التحصيلية لا سيما في اختبار المقال (essay) يمكن تقدير درجتها على الحكم الذاتي للمصحح مما يمكن أن تصحح إجابات المفحوصين بواسطة أثنين أو أكثر من الحكام أو الخبراء، ومن ثم أيجاد معامل ثبات المصححين على نفس المجموعة. بيد أنه نادراً ما تحتاج الاختبارات الموضوعية

إلى حساب معامل ثبات المصححين (Scorer Reliability) وذلك يعود لقلة أخطاء التصحيح فيها ولسهولتها في إعطاء الدرجة وتقدير عدد وحجم الأخطاء. إن معامل ثبات المصححين في هذه الحالة هو (معامل الارتباط) بين مجموعتي الدرجات التي يحصل عليها المفحوصين في العينة التي جاءت من تصحيح الباحثين المستقلين أو المصححين كل على انفراد، آلا أن إيجاد معامل ثبات المصححين لا يعني عدم الحاجة إلى معامل ثبات المجيب أو معامل ثبات الاختبار، بل ينبغي إيجاد معامل ثبات المجيب والاختبار إضافة إلى معامل ثبات المصححين، حيث أن معامل ثبات المصححين يشير إلى أي مدى يتفق اثنان أو أكثر في تقدير درجات نفس المجموعة على نفس الاختبار، بينما يشير معامل ثبات الممتحن (المجيب) إلى مدى تناسق أداء المجيبين على نفس مجموعة الأسئلة أو الفقرات، في حين يشير معامل ثبات الاختبار إلى مدى تشابه أداء الممتحنين على الفقرات أو الأسئلة المختلفة، ويفضل استخدام أسلوبTest-Retest في إيجاد معامل ثبات الممتحنين (المجيبين).

obeikan.com