الفصل الخامس
تقنين الاختبارات

تقنين الاختبار :-

يعرف التقنين بأنه رسم صورة شاملة لجميع خطوات الاختبار من حيث الفقرات ، والتعليمات والتصحيح وتفسير نتائج الاختبار، وتحديد المكان والزمان المناسبين ، وتحديد المصطلحات المناسبة التي تكون عامة لجميع المفحوصين ، (هناك بعض الاستثناءات توجد للطلاب المعوقين) *(Bracey , 2000 , 13) .*

خطوات التقنين :-

أولاً : الصدق :-

يشير مفهوم الصدق *Validity* إلي مدي قياس الاختبار لما يدعي أو يفترض أنه يقيسه *(Treece & Treece , 1977)* ولا يمكن قياس صدق الاختبارات التربوية أو النفسية بصورة مباشرة ، لكن يمكن الحكم علي صدق الاختبار علي أساس قاعدة عريضة من المعلومات تهدف إلي تحقيق عملية الصدق (فاروق عبد الفتاح ٢٠٠٧ ، ١٩٩) .

- **طرق حساب الصدق :-**

توجد عدة طرق لحساب الصدق هي : صدق المحتوي ، صدق المحك ، صدق التكوين . ويمكن تناول هذه الطرق علي النحو التالي :-

obeikandl.com

١- صدق المحتوي *Content Validity* :-

يطلق علي صدق المحتوي أسماء عدة منها الصدق السطحي أو الظاهري ، وصدق المحكمين والذي يجب ألا يقل عددهم عن خمسة ويكونوا من المتخصصين في المجال وفي القياس النفسي ، ولا تقل درجة الاتفاق علي كل بند من البنود عن ٨٠٪ (صلاح مراد أمين علي ، ٢٠٠٥ ، ٣٥١) ، وصدق المضمون ، والصدق المنهاجي ، والصدق المنطقي .

يشير ألين وين *Allen & Yen, 1979* إلي أن دراسة صدق المحتوي يتم من خلال عملية التحليل المنطقي لمحتوي أداة القياس اعتماداً علي حكم فردي (موسي النبهان ٢٠٠٤ ، ٢٧٤) ويركز الحكم علي درجة تمثيل البنود للمكونات الأساسية للسمة .

فإذا احتوي الاختبار علي عدد قليل من البنود وحذف بعض الجوانب الهامة من مقياس الرياضيات علي سبيل المثال أو تضمن موضوعات غير مناسبة ، فإن صدق المحتوي سينخفض . ومعظم الاختبارات التحصيل التجارية تم مراجعتها من قبل خبراء في المجال لتحديد ما إذا كان صدق المحتوي مناسب أو غير مناسب: *(Witt et al ., 1994 , 105)*.

وصدق المحتوي أو الظاهري ليس صدقاً بالمعني الحرفي والعلمي لكلمة الصدق . ويدل علي ما يبدو أن الاختبار يقيسه. بمعني أنه يشير إلي كيف يبدو الاختبار مناسباً للغرض الذي وضع من أجله (محمد شحاته ، ٢٠٠٨ ، ١١٧) .

obeikandl.com

وعلي الرغم من أن مفهوم صدق المحتوي مفهوم قديم إلا أن تطبيقاته ليست كثيرة ، وهذه الطريقة لا يسهل استخدامها بالاعتماد علي التكميم ، ومن هنا فقد استخدمت بطريقة عشوائية غير منهجية (جابر عبد الحميد ، ١٩٨٣- ب ، ٢٢١) .

ويُعتبر صدق المحتوي أسلوباً مناسباً في تقويم اختبارات التحصيل فهو يسمح لنا بالحصول علي إجابة لسؤالين رئيسين حول صدق الاختبارات المهنية والتحصيلية :-

(أ) هل يغطي الاختبار عينة ممثلة لمهارات ومعرفة محددة ؟

(ب) هل الأداء علي الاختبار محرر بدرجة كافية من تأثير المتغيرات غير المناسبة؟

ويعتمد صدق المحتوي علي إجراءات إحصائية أهمها :-

١- تحديد أي البنود يميز أفضل تمييز بين ذوي المستويات العليا والدنيا في الأداء.

٢- تحديد النسب المئوية لإجابة كل بند إجابة صحيحة .

٣- تحديد الدلالة الإحصائية لزيادة متوسط الدرجات من صف دراسي لآخر.

٤- تحديد معامل ارتباط كل بند بالتقدم التعليمي والتحصيل المدرسي عامة (الدرجات المدرسية) (عطية هنا ، محمد هنا ، ١٩٧٣ ، ١٨٢) .

ومن ناحية أخري فإن صدق المحتوي لا يناسب اختبارات الاستعدادات واختبارات الشخصية وقد تكون في الواقع مضللة (إسماعيل الفقي ، ٢٠٠٥ ، ٧٧).

وهذا النوع من الصدق في عدة مواضع يكون ضروري ، فهو يكون دافع للمفحوص إلي رؤية الاختبار علي أنه مناسب ، كما أنه يكون مفيد لكسب ثقة العميل بالاختبار *(Kaplan & Saccuzzo , 2001 , 133)* .

٢- صدق الارتباط بالمحك(الصدق الامبريقي *Related Validity–Criterion* :-

هناك أنواع متعددة من المحكات مثل التحصيل الأكاديمي العام وهو أشهر المحكات وخاصة بالنسبة لاختبارات الذكاء ، والأداء في برنامج معين وهذا المحك أكثر ارتباطاً بحساب صدق بطاريات الاستعدادات الفارقة ، واختبارات القدرات الخاصة ومقاييس بعض سمات الشخصية ، وهو يعتمد علي مؤشرات أداء المفحوصين في برنامج معين من التدريب المتخصص أو الأداء علي العمل نفسه وهو أكثر المحكات صلاحية لأنه يعتمد علي سجلات تتبعية للأداء الواقعي والفعلي للأفراد في العمل ، أو المجموعات المتضادة وهو من نوع المحكات المركبة التي تدل علي الآثار المتراكمة وغير المضبوطة للحياة اليومية فيمكن أن نقارن بين درجات الذكاء للمتخلفين عقلياً المقيمين بالمؤسسات بالدرجات التي يحصل عليها تلاميذ مدرسة عادية(فؤاد أبو حطب وآخرون ، ١٩٨٧ ، ١٥١-١٥٣) .

ومن أهم طرائق حساب صدق الارتباط بالمحك الصدق التلازمي *ConcurrentValidity* والذي يتناول تحديد العلاقة الارتباطية في حالة تطبيق الاختبار والمحك *(Kaplan & Saccuzzo , 2001 , 135)* وفي هذا النوع يتم مقارنة أداء الأفراد علي المقياس موضع الدراسة بأداء نفس الأفراد علي مقياس آخر قد سبق التحقق من صدقه والذي يسمي بالمحك الخارجي . والمحك الخارجي هو مقياس أو اختبار أخر يقيس نفس السمة التي يقيسها المقياس الجديد فعلي سبيل المثال إذا صممت أداة لقياس القدرة الأكاديمية فإن درجات الطلاب علي هذه الأداة يمكن مقارنتها بالمعدل التراكمي لهؤلاء الطلاب (وهو ما يُعرف بالمحك الخارجي) فإذا كانت الأداة تقيس بالفعل القدرة الأكاديمية فعليه يجب أن تكون نتائج الطلاب في الأداة متناسبة مع معدلهم التراكمي (صلاح مراد ، فوزية هادي ٢٠٠٢ ١٨٦) .

وتتمثل أهم مشكلات الصدق التلازمي في صعوبة تحديد علاقات السبب والأثر لأنه يدل بالفعل علي علاقة اقتران أكثر من دلالته علي علاقة علية .

وللصدق التلازمي استخداماته في أغراض معينة فهو أكثر ملائمة للاختبارات التشخيصية والتي نلجأ إليها في حل المشكلات الراهنة أكثر من مشكلات المستقبل (فؤاد أبو حطب وآخرون ، ١٩٨٧ ، ١٥٠) .

والصدق التنبؤي *Predicative Validity* والذي يشير إلي قدرة الاختبار وفاعليته في التنبؤ بنتيجة معينة في المستقبل ، ولهذا يقوم علي المقارنة بين درجات المفحوصين علي الاختبار ودرجاتهم في مقياس آخر مباشر للأداء اللاحق للفرد . ويستند علي أساس حساب القيمة التنبؤية للاختبار ، فنتائج التوجيهية تنبئ بالنجاح في كليات المجتمع أو الدراسة الجامعية (مصطفي القمش وآخرون ، ٢٠٠٠ ، ١١١) .

٣- صدق التكوين الفرضي *Construct Validity* :-

لتحديد ما إذا كان الاختبار يتضمن صدق التكوين يجب أن نعتمد علي النظرية التي تقف خلف التحليل الإحصائي أو التكويني لدرجات الاختبار ، علي سبيل المثال النظرية النفسية التي تشير إلي وجود ارتباط دال بين القلق ودرجات الامتحان لطلاب الجامعة ، هذه النظرية تنبئ أن الأفراد ذات المستوي المتوسط من القلق يؤدون بطريقة أفضل من الدرجات المنخفضة جداً أو المرتفعة جداً فالأولي تنشط وتدفع الطلاب إلي التحصيل ، والثانية تعوق الأداء ، فالاختبارات التي تقيس القلق تكون صادقة مع هذا التنبؤ إذا الطلاب الذين حصلوا علي

درجات مرتفعة ومتوسطة ومنخفضة من القلق سوف يؤدون وفقاً لهذه الفرضية *(Witt et al ., 1994 , 109)* .

ومن أكثر الطرق شيوعاً للحصول علي شواهد علي استخدام طريقة الصدق التكويني :-

- مقارنة درجات مجموعة من الأفراد قبل وبعد معالجة معينة كما يحدث حين نتنبأ بأن درجات التلاميذ علي مقياس الاتجاهات النفسية سوف ترتفع بعد تعليمهم علي يد معلمين ومربين ذوي اتجاهات إنسانية .
- مقارنة درجات جماعات معروفة كما يحدث حين نتنبأ بوجود فروق كبيرة بين المواطنين الراشدين وبين المراهقين علي أداة جديدة لقياس ميول المراهقين (جابر عبد الحميد ، ١٩٨٣- ب ، ٢٢٤-٢٢٥) .
- طريقة المقارنات الطرفية *Comparison of Extreme Groups* وهذه الطريقة في حال الرغبة في تعرف مدي قدرة الاختبار علي التفرقة بين المرتفعين والمنخفضين في السمة المقاسة ، أو القدرة علي التمييز بين المستويات المختلفة للسمة. ونشير هنا إلي أن هذه القدرة علي التمييز لا تعد صدقاً ، وإنما تدل علي اتساع مدي درجات السمة المقاسة. وما نقصد بالمقارنات الطرفية هو استخدام محك آخر يصنف الأفراد إلي مستويات في السمة المرغوبة ، ثم تطبيق الاختبار علي هؤلاء الأفراد وبحث الفروق بين درجات الاختبار لمجموعات المستويات المختلفة ، وفي هذه الحالة تدل الفروق علي صدق تجريبي أو صدق المحك لدرجات الاختبار مثل صلاحية اختبار للذكاء في التمييز بين المتخلفين عقلياً والأسوياء(صلاح مراد ، أمين علي ، ٢٠٠٥، ٣٥٦-٣٥٧).

obeikandl.com

- الصدق التقاربي والصدق التمييزي *Convergent & Discriminant Validity* أشار كامبيل *Campbell , 1960* في دراسة تحليلية عن صدق التكوين الفرضي أنه لكي نظهر صدق التكوين فإن علينا أن نوضح أن الاختبار لا يرتبط فقط بدرجة عالية مع المتغيرات الأخرى التي ينبغي أن يرتبط بها نظرياً وإنما أيضاً أنه لا يرتبط بشكل جوهري مع المتغيرات التي ينبغي أن يختلف معها ، وفي مقال سابق وصف كامبل وفيسك (*Campbell & Fiske , 1959*) العملية الأولي علي أساس أنها الصدق التقاربي والثانية علي أنها الصدق التمييزي علي سبيل المثال الارتباط العالي بين اختبار الاكتئاب وآخر للميل العصابي يمكن أن يؤخذ مثالاً للصدق التقاربي بينما الارتباط المنخفض بين اختبار للاكتئاب واختبار للمرح يُعتبر مثالاً للصدق التمييزي (إسماعيل الفقي ، ٢٠٠٥ ، ١٠١-١٠٢) .

- الصدق العاملي *(Factorial Validity)* وفي هذا النوع من الصدق يستخدم التحليل العاملي في بيان إلي أي مدي يقيس الاختبار السمة أو الظاهرة التي وضع لقياسها وذلك بحساب درجة تشبع الاختبار بهذه السمة أو الظاهرة . ويتحقق ذلك بحساب معاملات ارتباط الاختبارات والموازين المختلفة ، ثم تحلل هذه العوامل عاملياً لمعرفة أقل عدد ممكن من العوامل تعتبر مسئولة عن هذا الارتباط سواء كانت هذه العوامل عامة أو طائفية (رمزية الغريب ، ١٩٨٨ ، ٦٨٢) .وبعد ذلك نبدأ عملية تفسير العامل الإحصائي الذي تم الحصول عليه ، فإذا كانت الاختبارات من نوع الأداء الأقصى وبالطبع نحن لا نعتمد في هذا التحليل علي مجرد تأمل مصفوفة الارتباط وإلا كانت العملية

صعبة ومعقدة وغير مؤكدة بل مستحيلة وخاصة ، إذا كانت بطارية الاختبارات الكبيرة ، ولذلك تستخدم الأساليب الإحصائية الأكثر دقة لتحديد العوامل المشتركة التي يمكن اعتبارها مسئولة عن معاملات الارتباط الناتجة (فؤاد أبو حطب وآخرون ، ١٩٨٧ ، ١٦٣) .

- الاتساق الداخلي *Internal Consistency* يتم حساب الاتساق الداخلي بمعاملات الارتباط بين الدرجة الكلية ودرجات مكونات بطارية الاختبار (أو علاقة درجات بنود الاختبار بالدرجة الكلية إذا كان يقيس شيئاً واحداً) وتدل معاملات الارتباط هذه علي أن المكونات أو البنود تقيس شيئاً مشتركاً مما يعني صدق البناء الداخلي ولكنه لا يعني صدق أو صلاحية الأداة لقياس السمة موضع الاهتمام. لذلك يجب عدم الاعتماد الكلي علي هذه الطريقة في محاولات التحقق من صدق أو صلاحية استخدام درجات الاختبار (صلاح مراد ، أمين علي ، ٢٠٠٥ ، ٣٥٧) .

٤- الصدق الذاتي *Intrisc Validity* :-

يعرف الصدق الذاتي بأنه صدق الدرجات التجريبية بالنسبة للدرجات الحقيقية التي خلصت من شوائب أخطاء الصدفة ، وبذلك تصبح الدرجات الحقيقية للاختبار هي الميزان أو المحك الذي تنسب إليه صدق الاختبار ، ولما كان ثبات الاختبار يؤسس علي ارتباط الدرجات الحقيقية للاختبار بنفسها إذا أعيد الاختبار علي نفس المجموعة التي أجري عليها في أول الأمر. لهذا كانت الصلة وثيقة بين الثبات والصدق الذاتي (رمزية الغريب ، ١٩٨٨ ، ٦٨٣) .

والعلاقة بين الصدق الذاتي والثبات تعبر عنهم هذه المعادلة :-

الصدق الذاتي = √ معامل الثبات

وقد وضع الباحثون درجات الصدق المقبولة كما يلي :

درجة ٧٠ لغايات الدراسة ، ٩٠ لغايات اتخاذ القرار والحكم .

٨٠- ٩٩ صدق عال.

٥٠-٧٩ موضع تساؤل.

٥٠ فما دون غير مقبول (عصام النمر، ٢٠٠٨ ، ٧٠) .

- العوامل المؤثرة في الصدق :-

هناك عدة من العوامل المؤثرة في الصدق وهي علي النحو التالي :-

عوامل في الاختبار نفسه ومن أهمها :-

١- المفردات والتراكيب الصعبة .
٢- البنود بين الصعبة جداً والسهلة جداً .
٣- البنود الموحية بالإجابة .
٤- الغموض حيث يؤدي إلي سوء التفسير.
٥- ترتيب البنود بصورة غير ملائمة (أمطانيوس ميخائيل ، ١٩٩٥ ، - ٢٧١- ٢٧٢) .
٦- وضوح الأهداف للاختبارات وبالتالي يمكن مقارنة ذلك بمادة الاختبار ومعرفة. درجة الصدق من خلال المعادلات الإحصائية (عصام النمر، ٢٠٠٦ ، ٥٩) .

obeikandl.com

٧- العوامل المتصلة بشروط الإجراء والتصحيح ومنها :

أ- الشروط البيئية المحيطة .

ب- الوقت المخصص للإجابة .

جـ- أخطاء التصحيح (أمطانيوس ميخائيل ، ١٩٩ ، - ٢٧١-٢٧٢) .

ثانياً : الثبات :-

- **مفهوم الثبات :-**

يشير الثبات إلي اتساق الدرجات التي تم الحصول عليها من جراء تطبيق أداة ما. أي مدي اتساق درجات المقياس إذا ما أعيد تطبيقه علي نفس الأفراد (صلاح مراد ، فوزية هادي ، ٢٠٠٢ ، ١٩١).

وثبات الاختبار مفهوم هام بالنسبة للخصائص القياسية لأي اختبار أو مقياس . فإذا كان الاختبار ثابتاً وكانت السمة المقاسة أو السلوك المقاس ثابتاً فسوف يحصل الفرد علي نفس الدرجة في مرات تطبيق الاختبار المتكررة . ويكون مقدار النقص في ثبات الاختبار مساوياً لمقدار اختلاف درجات الفرد في التطبيق المتكرر لهذا الاختبار. وفي القياسات التربوية والنفسية نكون في حاجة ماسة إلي مقاييس واختبارات ذات درجة عالية من الثبات قدر الإمكان(فاروق عبد الفتاح ٢٠٠٧ ، ١٦٥) .

تصنف وسائل تقرير الثبات في فئتين :-

١- الثبات النسبي : يذكر علي أنه معامل ارتباط ، أو معامل ثبات ، ويدل هذا المعامل الإحصائي علي المدى الذي تكون فيه نتائج تطبيق الاختبار بصورتين متكافئتين علي مجموعة من الأفراد متسقة نسبياً . وفي بعض الأحيان يقصد بالثبات النسبي تحليل التباين .

٢- الثبات المطلق : ويذكر علي أنه الخطأ المعياري للمقياس ومقدار انحراف الدرجات الناتجة عن حقيقتها (عطية هنا ، محمد هنا ، ١٩٧٣ ، ١٨٧) .

• طرق حساب الثبات :-

توجد عدة طرق لحساب ثبات الاختبار يمكن تناولها بالتفصيل علي النحو التالي :-

١- معامل الاستقرار *Stability Coefficient* :-

غالباً ما يطلق علي معامل الاستقرار اسم تقدير الثبات عن طريق إعادة الاختبار *A test –Retest Estimate of Reliability* ويمكن الحصول علي هذا النوع من الثبات عن طريق إعطاء الاختبار للطلبة ثم إعادة إجرائه لنفس الطلبة فيما بعد بنفس الطريقة ، وتحت نفس الظروف تقريباً ، ثم حساب معاملات الارتباط بين مجموعتي الدرجات ، وهنا نحن نقيس استقرار الدرجات عبر عدة أداءات للطلبة في الاختبار (قاسم علي الصراف ، ٢٠٠٢ ١٩١) *(Witt et al ., 1994 , 96)* .

وهذه الطريقة من أفضل الطرق المستخدمة في حساب معامل ثبات الاختبارات غير الموقوتة (رمزية الغريب ، ١٩٨٨ ، ٦٧١) وتصلح هذه الطريقة أيضاً للاختبارات الموقوتة والتي تعتمد إلي حد كبير علي السرعة ، ولا تصلح هذه الطريقة لحساب ثبات الاختبارات التي تقيس التذكر أو ترتبط ارتباطاً مباشراً بهذه العملية العقلية ، وذلك لتأثر عملية التذكر تأثراً مباشراً بالفاصل الزمني بين مرتي إجراء الاختبار ، لذلك يجب ألا يتجاوز الفاصل الزمني بين التطبيقين أسابيع قليلة بالنسبة للأطفال ، وألا يتجاوز ستة أشهر بالنسبة للكبار البالغين (أحمد غنيم ، نصر محمود ، ٢٠٠٠ ، ٢٥٩) .

وهذا النوع يكون ذا قيمة فقط عندما تكون لقياس سمات أو الخصائص التي لا تتغير عبر الزمن ، فعلي سبيل المثال نحن غالباً نفترض أن اختبارات الذكاء

تقيس قدرة عامة ثابتة أو مستقرة ، فإذا تم تطبيق المقياس في مرتين وحصلنا علي درجات مختلفة نحن نستنتج أن هناك قصور في التطابق يرجع إلي وجود خطأ عشوائي في القياس غالباً نحن لا نفترض أن الشخص ذكي وأقل ذكاءاً في الفترة التي بين التطبيقين :*(Kaplan & Saccuzzo , 2001 , 106)* .

عيوب طريقة إعادة الاختبار :-

١- إن الطلبة المفحوصين قد يتعلمون من الاختبار في المرة الأولي ويألفونه مما يجعل نتائجهم المتحققة عليهم في المرة الثانية أفضل .

٢- تذكر المفحوصون بعضاً من أجزاء الاختبار إذا كانت الفترة الفاصلة بين التطبيقين قصيرة .

٣- إن العوامل النفسية لدي الطالب المفحوص تختلف عند إجراء الامتحان في المرة الأولي وإجرائه في المرة الثانية (إبراهيم محمد ، ٢٠٠٧ ، ١٠١) .

٤- إذا كانت المدة الفاصلة بين التطبيقين طويلة فإن هناك الكثير من العوامل الأخرى من الممكن أن تؤثر في الدرجات فعلي سبيل المثال إذا كان التطبيق علي أطفال في الرابعة والخامسة من العمر وكان معامل الارتباط بين التطبيقين ٠,٤٣ فهناك عدة من التفسيرات المحتملة لهذا الارتباط المنخفض:-

- أن ثبات الاختبار متدني.
- أن التغير حدث في خصائص الأطفال في ٤ و ٥ سنوات(تأثير عامل النضج) .
- أو حدث بعض الدمج بين الثبات المنخفض والتغير في الأطفال يكونان مسئولين عن هذا الارتباط المنخفض.

obeikandl.com

- أكثر من هذا معظم التقويمات التي تتبع إعادة تطبيق الاختبار لا ترجح أحد التفسيرات علي الأخرى *(Kaplan & Saccuzzo , 2001 ,107)* .

٢- الصور المتكافئة(معامل التكافؤ *Equivalence Forms*) :-

يفترض في هذه الطريقة أنه يمكن إنشاء صورتين متكافئتين من الاختبار. وتطبق صورتي الاختبار علي نفس المجموعة من الأفراد في نفس اليوم ، بفاصل زمني قصير (عدة ساعات) ، *(Kaplan & Saccuzzo , 2001 , 108)* .

ومعامل الثبات في هذه الحالة هو معامل الارتباط بين درجات مجموعة من الأفراد علي هاتين الصورتين المتكافئتين ، ويمكن أن نسمي احدي الصورتين الصورة (أ) والثانية الصورة (ب) (محمد شحاته ، ٢٠٠٨ ، ٨٥) .

ويشترط التكافؤ أن يتحقق ما يلي : -

١- أن تكون المجالات والموضوعات التي يقيسها المقياس واحدة .
٢- تساوي نسبة البنود التي تخص كل موضوع من الموضوعات .
٣- تساوي المتوسط الحسابي والانحراف المعياري لدرجات الأفراد علي كل من الصورتين .
٤- تشابه طريقة صياغة البنود .
٥- تماثل مستوي صعوبة البنود .
٦- تساوي طول الصورتين وطريقة إجراؤهما وتصحيحهما وتوقتيهما (صلاح مراد ، فوزية هادي ، ٢٠٠٢ ، ١٩٣) .

وتعالج طريقة الصور المتكافئة بعض الانتقادات الموجهة لطريقة إعادة الاختبار من حيث تذكر أسئلة الاختبار ، وعدم توفر نفس ظروف التطبيق الأول ،

obeikandl.com

والنمو العقلي والجسمي والاجتماعي لأفراد العينة . وعموماً فإن استخدام طريقة إعادة الاختبار أو الصور المتكافئة يتوقف علي الأغراض التي من أجلها يطبق الاختبار (أحمد غنيم ، نصر محمود ، ٢٠٠٠ ، ٢٦٠) .

كما أنها طريقة مفضلة في تقدير الثبات وتستخدم كثيراً في اختبارات الاستعدادات والقدرات .

عيوب طريقة الصور المتكافئة :-

١- التعب والملل اللذان قد يحدثا عند التلميذ فينخفض معامل الثبات (مصطفي القمش وآخرون ، ٢٠٠٠، ١١٤) .

٢- صعوبة إيجاد صورتين متكافئتين لنفس الاختبار *(Witt et al ., 1994 , 98)*

٣- كما نجد بعض الصعوبات نواجهها في تهيئة بيئة طبيعية مماثلة في فترتين التطبيق ، فضلاً عن أنها مكلفة مادياً .

٣- معاملي التكافؤ والاستقرار *Coefficient of Equivalence and Stability* :-

هناك حالات نضطر فيها إلي الاهتمام بالتنبؤ والاستنتاج علي المدى البعيد لمجال من مجالات المعرفة علي سبيل المثال : قياس البناءات العقلية كالذكاء والإبداع والميول الموسيقية لا يعتمد بشكل عام علي مجموعة من الأسئلة ، وإذا اعتمد علي ذلك ، فإن البناء العقلي لا يثير كثيراً اهتمامنا نحن نحب أن نعرف ما إذا كانت مجموعة من الأسئلة المختلفة ، ولكنها متشابهة ، إذا ما استخدمت في أوقات مختلفة ، تعطي نتائج متشابهة أم لا (قاسم علي الصراف ، ٢٠٠٢ ، ١٩٢) .

وفي هذه الحالة فإن معامل الثبات يمكن الحصول عليه عن طريق إعطاء الصورة الأولي من الاختبار، وبعد فترة من الوقت يتم تحديدها تعطي الصورة

الأخرى المكافئة للصورة الأولي ثم نقارن النتائج عن طريق حساب معامل الارتباط بينهما (أحمد غنيم ، نصر محمود ، ٢٠٠٠ ، ٢٦١) .

٤- معامل الاتساق الداخلي(طريقة التجزئة النصفية):-

Intern Consistenc Coefficient

لا تتطلب طريقة التجزئة النصفية *Spilt –Half Method* إعادة تطبيق الاختبار أو وجود صورتين متكافئتين ، وإنما تقوم علي تطبيق الاختبار مرة واحدة وتقسيمه إلي نصفين(فردي- زوجي) ويتم حساب معاملات الارتباط بين نصفي الاختبار:*(Kaplan & . Saccuzzo , 2001 , 108)* .

مزايا هذه الطريقة :-

من مزايا هذه الطريقة أنها تتخلص من الكثير من العيوب السابقة فهناك تشابه في ظروف التطبيق ، كما أنها لا تتأثر بعامل النضج ، أو الممارسة ، كما أنها توفر في الجهد والمال .

عيوب هذه الطريقة :-

أ- أنها لا تسمح باختبار الفرد إلا مرة واحدة فإذا تصادف وجود ظروف معرقلة للتحصيل في الاختبار مثل الحالة الصحية أو قلة الدافعية أو أي سبب خارجي أخر تأثرت بها النتائج (رمزية الغريب ، ١٩٨٨ ، ٦٥٧-٦٥٨).

ب- كيف نجزئ الاختبار للحصول علي أفضل قسمين للمقارنة وخاصة إذا علمنا أن معظم الاختبارات لا يمكن تجزئتها إلي نصفين صالحين للمقارنة بينهما بمجرد تتبع التسلسل الرقمي للأسئلة ، وذلك بسبب الفروق في طبيعة

openkandil.com

الأسئلة ومستوي صعوبتها بالإضافة إلي الآثار التراكمية للحماس في الأداء أو العمل والممارسة والتعب والملل (فؤاد أبو حطب وآخرون ، ١٩٨٧ ، ١١٥) .

وهناك الكثير من المعاولات لحساب معامل الثبات بطريقة التجزئة النصفية :-

" معادلة " سيبرمان وبراون"

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

حيث أن ر = معامل الثبات الذي نريد الحصول عليه .

ر١ = معامل الثبات الذي حصلنا عليه.

ن = عدد مرات إطالة الاختبار أو اختصاره .

ويمكن اختصار المعاولة حينما تطبق علي معامل الثبات المقرر بطريقة التجزئة النصفية لأنها تتضمن إطالة الاختبار مرتين فتكون علي النحو التالي :

$$ر = \frac{٢\ ر_{١}}{١ + ر_{١}}$$

وهي تستخدم في حالة النصفين المتكافئين للاختبار ولا تصلح لحساب ثبات الاختبار إذا انقسم إلي جزأين غير متكافئين وخاصة عندما تختلف القيم العددية للتباين اختلافاً كبيراً ، وذلك لأن البرهان الرياضي لمعادلة التنبؤ هذه يفرض تساوي (تكافؤ) الجزأين في البناء الإحصائي (أحمد غنيم ، نصر محمود ، ٢٠٠٠ ، ٢٦٢) .

ولا تصلح هذه الطريقة لحساب الاختبارات الموقوتة التي تعتمد اعتماداً كبيراً علي سرعة الاستجابات لأن كثرة الأسئلة المتروكة في آخر كل اختبار تؤثر

obeikandl.com

علي الارتباط بين الجزئيين ويتغير بذلك معامل الثبات(فؤاد البهي ، ١٩٧٩، ٥٢٦).

معادلة رولون المختصرة للتجزئة النصفية :-

وتهدف هذه الطريقة إلي تبسيط معادلة سبيرمان براون وذلك بحساب تباين فروق درجات النصفين وحساب تباين درجات الاختبار.

والقانون هو :

$$ر_{11} = 1 - \frac{ع^2_{(1-2)}}{ع^2}$$

حيث أن :

ر١١= معامل الثبات .

ع (١-٢)٢= تباين الفرق بين درجات النصف الأول ودرجات النصف الثاني للاختبار.

ع٢= تباين درجات الاختبار كله.

٥- طريقة التباين :-

أ- معامل كيودور - ريتشاردسون(معامل التجانس) *Homogeneity Coefficient* :-

تعتمد هذه الطريقة أيضاً علي تطبيق الاختبار في مرة واحدة . ويمكن باستخدامها التحقق من تجانس جميع مفردات الاختبار، وهذا التجانس يتأثر بكل من معاينة المحتوي كما في طريقة التجزئة النصفية ، وعدم اتساق النطاق السلوكي الذي تستمد منه عينة مفردات الاختبار. فكلما زاد هذا الاتساق دل ذلك علي تجانس مفردات الاختبار. فالاختبار الذي يشتمل علي مفردات تقيس عملية الجمع البسيط فقط ، يكون أكثر اتساقاً أو تجانساً من الاختبار الذي يقيس عمليتي الجمع والطرح . وتتغلب هذه الطريقة علي مشكلة تجزئة الاختبار إلي نصفين

obeikandl.com

متكافئين ، حيث أنها تعتمد علي إيجاد قيمة تقديرية لمتوسط قيم معاملات الارتباط بين كل من نصفي الاختبار إذا تم تجزئته بجميع الطرق الممكنة دون أن نلجأ إلي هذه التجزئة فعليا(صلاح علام ، ٢٠٠٦ ، ٩٨) .

تصلح هذه الطريقة للاختبارات التي تتم الإجابة علي أسئلتها بالثنائية ، من نوع الكل أو لا شيء مثل صواب وخطأ ، نعم-لا ، موافق - معارض وغيرها وهي لا تناسب الاختبارات التي تعتمد علي السرعة . وتعد طريقة التجزئة النصفية حالة خاصة من هذه الطريقة .

ومعادلة كيودور - ريتشاردسون تتم علي النحو التالي:-

$$ر = \left(\frac{ن}{ن-١}\right) = \frac{ع٢ - مج\ ص\ خ}{ع٢}$$

حيث أن ر = معامل ثبات الاختبار.

ن = عدد الأسئلة التي يتكون منها الاختبار.

ع٢ = تباين الاختبار ككل .

مج ص خ = يحسب بأعداد نسبة الأفراد الذين أصابوا في إجاباتهم علي كل سؤال ونسبة الذين أخطأوا فيه ونضرب النسبتين في بعضهما ، ونجمع حاصل الضرب بالنسبة لجميع الأسئلة (فؤاد أبو حطب وآخرون ، ١٩٨٧ ، ١١٨) .

ب- معامل ألفا كرونباخ *Coefficient Alpha* :-

وهي معادلة كرونباخ وتستخدم لقياس الاتساق الداخلي للاختبار عن طريق تحليل المفردات الداخلية وقياس مدي اتساقها بعضها مع البعض الأخر.

وهي طريقة تناسب اختبارات القوة التي تشتمل علي مفردات متعددة البدائل وذلك من خلال المعادلة التالية :-

$$\text{ر} = \frac{\text{ن}}{\text{ن} - 1} \left(1 - \frac{\text{مج ع}^2 \text{أ}}{\text{مج ع}^2 \text{ك}} \right) .$$

حيث: ك = عدد مفردات الاختبار.

مج ع٢ك = مجموع التباين لكل مفردة من مفردات الاختبار.

مج ع٢ك = مجموع التباين الكلي للأفراد .

وتستخدم معامل ألفا في حالة الاختبارات المدرجة بطريقة ليكرت(بطرس حافظ،٢٠٠٣ ، ٦٥) .

ومن الجدير بالذكر أن صيغتي كيودور- ريتشاردسون وكرونباخ تستندان إلي أن اتساق استجابات الأفراد عبر مفردات الاختبار يمكن الاعتماد عليه في تقدير معامل الثبات للدرجات الكلية في الاختبار. فإذا كانت مفردات الاختبار متجانسة ، أي تقيس سمة أو بُعد واحد ، فإن كلاً من الصيغتين تؤدي إلي نتيجة قريبة من الدرجة التي نحصل عليها من تطبيق طريقة التجزئة النصفية . أما إذا كانت المفردات غير متجانسة فإن قيمة معامل الثبات الناتجة عن تطبيق كل من الصيغتين سوف تكون أقل من القيمة الناتجة عن التجزئة(صلاح علام ، ٢٠٠٦ ، ١٠١) .

٦- ثبات المصححين :-

يمكن حساب الثبات عن طريق تقديرات المحكمين هناك حالات يلجأ إليها الباحثون إلي مجموعة من المحكمين للحصول علي مؤشر لثبات مقاييسهم واختباراتهم ، كما في حالة قياس الشخصية الذي يعتمد علي الملاحظة والمقابلة الشخصية ، أو مقياس التقدير أو الطرق الإسقاطية أو قياس التحصيل الدراسي .

وفي هذه الحالة يتم تحديد النسبة المئوية للاتفاق بين المحكمين (قاسم علي الصراف ، ٢٠٠٢ ، ١٩٥) .

ويطلق علي هذه الطريقة ثبات الاتفاق بين المصححين *Inter – Scorer Reliability* ويكون حساب معامل بيتا المصححين يُعتبر بسيطاً حيث تجري تصحيح عينة من الاختبارات بشكل مستقل بواسطة مصححين أو أكثر ثم يحسب معامل الارتباطات بين أزواج المصححين . وفي المعتاد أن يثبت معدو الاختبار في دليل الاختبار بياناً بالخبرة والتدريب المطلوبين في الفاحصين ثم تثبت معامل ثبات المصححين (إسماعيل الفقي - ٢٠٠٥ ، ٤٩) .

وبصورة عامة يجب ألا يقل معامل الثبات عن ٨٠ ويعدل في المقياس ويعاد تطبيقه إذ قل عن ذلك (عصام النمر ، ٢٠٠٨ ، ٨٨) .

اختيار الطريقة المناسبة لحساب ثبات الاختبار :-

مصادر الخطا المحتمل حدوثها في طرق تقدير الثبات :

المصادر المحتملة للخطأ	طرق تقدير الثبات				
	الاستقرار	التكافؤ	الاستقرار والتكافؤ	الاتساق الداخلي	ثبات التصحيح
عدم استقرار السمة	"		"		
خطأ معاينة		"	"	"	
إدارة الاختبار	"	"	"		
خطأ عشوائي ضمن الاختبار	"	"	"	"	"
خطأ التصحيح					"

يتضح من هذا الجدول أن الخطأ العشوائي يمكن أن يكون ذا تأثير علي كافة طرق تقدير الثبات (موسي النبهان ، ٢٠٠٤ ، ٢٥٥) .

العوامل التي تؤثر علي ثبات الاختبار :-

- **طول الاختبار :-**

زيادة عدد فقرات الاختبار مما يقلل من التباين لشمول عينة من السلوك أكبر يزداد معامل ثبات الاختبار مع زيادة عدد الأسئلة التي تكونه . ومن مساوئ هذه الطريقة أنها قد تؤدي إلي تعب المفحوص والملل ، وخاصة في اختبارات التحصيل ولذلك يمكن التحكم بمعامل الثبات بزيادة عدد فقراته (عصام النمر ، ٢٠٠٦ ، ٦٤) .

- **تباين الخطا :-**

يؤثر في ثبات الاختبار أي مصدر لتباين الخطأ يكون فعالاً في الموقف الاختباري وعلي الباحث النفسي أن يجرب كل سبيل ممكن للتأكد من ذلك (فؤاد أبو حطب وآخرون ١٩٨٧ ، ١٢٨) .

- **تجانس المجموعة :-**

يقل ثبات الاختبار بزيادة تجانس المجموعة المطبق عليها ، ويرتفع الثبات بزيادة تباين درجات أفرادها عليها (إسماعيل الفقي ، ٢٠٠٥ ، ٦٦) .

- التخمين في الإجابة علي بنود المقياس : صعوبة البنود وغموض التعليمات تجعل الأفراد القليلي الاستعداد يعتمدون في إجاباتهم علي التخمين مما يقلل من ثبات الأداة.
- استقلالية بنود الأداة بعضها عن بعض : حيث أن الارتباط المرتفع بين بنود الأداة يقلل من معامل ثباته ، حيث تؤدي هذه الخاصية إلي حالة خفض عدد البنود .
- زمن المقياس : حيث أن معامل الثبات ينخفض إذا كان الزمن أكبر من اللازم لذلك يجب أن تكون الأدوات موقوتة بفترة زمنية محددة(صلاح مراد فوزية هادي ، ٢٠٠٢ ١٩٩) .

ويشير ثورندايك Thorndike , 1949 إلي مجموعة من النقاط الهامة التي تؤثر في ثبات الاختبار:

(أ) خصائص عامة ودائمة في الأفراد :-

- المهارات العامة (علي سبيل المثال :القراءة) . القدرة العامة علي فهم التعليمات وفنيات تعاطي الاختبار.
- القدرة علي حل المشكلات لنمط أو نوع عام في الاختبار. الاتجاهات وردود الفعل الانفعالية والعادات التي عموماً تؤثر في الأداء مثل موقف الاختبار (علي سبيل المثال : الثقة بالنفس) .

(ب) الخصائص النوعية والدائمة للأفراد :-

- المعرفة والمهارات المطلوبة لمشكلات نوعية أو محددة في الاختبار.
- الاتجاهات وردود الفعل الانفعالية والعادات المرتبطة بمثيرات محددة أو نوعية بالاختبار (علي سبيل المثال : الخوف من الأماكن المرتفعة والتي يستحضرها المفحوص عندما يكون السؤال متعلق بالخوف من الأماكن المرتفعة) .

(ج) الخصائص العامة والمؤقتة :-

- الحالة الصحية ، والتعب ، والتوتر الانفعالي . الدافعية ، التجاوب مع الفاحص .
- تأثيرات الحرارة والضوء .....الخ .

(د)الخصائص النوعية والمؤقتة :-

- التغيرات التي تطرأ علي المفحوص في حالة التعب أو الدافعية التي تتطور من خلال اختبار معين (علي سبيل المثال : عدم التشجيع في حالة الفشل في بنود محددة من الاختبار). التقلب أو التذبذب الناتج عن الانتباه ، وعدم التآزر، ومعايير الحكم.

- التقلب أو التذبذب في الذاكرة لحقائق معينة.مستوي ممارسة المهارات أو المعرفة المطلوبة لهذا الاختبار المحدد أو النوعي .
- الحالات الانفعالية المؤقتة ، وقوة العادات *(in :Witt et al ., 1994 , 104)* .

• العلاقة بين الصدق والثبات :-

أوضح بوليت وهنجلر *(Polit & Hungler , 1987)* أن الثبات يكون ضروري ولكنه ليس كافياً للصدق . فيمثل الثبات شرطاً ضرورياً وغير كاف للصدق ، ومن الواضح أن المقياس الذي يعطي نتائج غير متسقة لا يمكن أن يعطي نتائج صادقة علي حين أن المقياس الذي يعطي نتائج غير صادقة لابد أن يعطي نتائج متسقة ، ويتمتع بصفة الثبات وبطبيعة الحال فإن الزيادة في معامل الثبات قد تؤدي إلي الزيادة في معامل الصدق ولكنها لا تحتم هذه الزيادة . فقد تنتج الزيادة في معامل الثبات عن ازدياد التباين الناتج عن الخطأ (أمطانيوس ميخائيل ، ١٩٩٥ ، ٢٧١-٢٧٢) .

ثالثاً :المعايير *Norms* :-

هي نوع من الموازين (المحكات) التي تستخدم في تفسير الدرجات الخام التي يحصل عليها الطالب . كما أن إعدادها يعتمد أيضاً علي الدرجات الخام لعينة معيارية ممثلة للمجتمع الذي أعدت له الأداة أو الاختبار(صلاح مراد ، أمين علي ، ٢٠٠٥ ، ٣٦٩) .

وهناك مجموعة من العوامل ذات الأهمية الخاصة بمعايير الاختبارات ، يمكن تحريرها في النقاط التالية :-

•معايير العمر : *Age Norms* :-

عيوب هذه الطريقة :-

١- النمو العقلي أو التحصيلي لا يساير النمو الزمني ومن هنا فإن النسبة لا تظل ثابتة كما يفترض ذلك .

٢- أن الذكاء لا يستمر في النمو طوال حياة الإنسان وإنما تقف عند سن معينة (عند حوالي ١٥-١٧ سنة) ولذلك فمهما تقدم عمر الفرد فإننا نفترض حداً ثابتاً لنموه الزمني وهو السن الذي يتوقف عنده الذكاء .

٣- النمو التحصيلي لا يستمر في النمو طوال العام بمعدلات ثابتة ولكنه يختلف بين مادة وأخري .

٤- تتأثر معايير العمر بعوامل بيئية ومدرسية متعددة (محمود عبد الحليم ١٩٨٠ ١١٣-١١٤) .

- **معايير الفرقة الدراسية** *Grade Norms* :-

معيار الفرقة الدراسية هو متوسط أداء فرقة دراسية معينة علي اختبار معين . وهنا نقارن درجة الشخص الذي نختبره بمتوسط الفرقة الدراسية التي ينتمي إليها. ومثال ذلك ، اختبار تحصيلي مقنن لمقرر الجغرافيا للمستوي الأول ثانوي ، هذا الاختبار يحتوي ١٠٠ سؤال ، والمتوسط الحسابي لدرجات عينة من طلاب المستوي الأول ثانوي ٦٠ ، وعلي ذلك فإن الطالب الذي تزيد درجته عن ٦٠ يكون أعلي من المتوسط ، أما الطالب الذي تقل درجته عن ٦٠ يكون أقل من المتوسط ، أما الذي يحصل علي ٦٠ فهو متوسط (محمد شحاته ، ٢٠٠٨ ، ٥٤) .

عيوب هذه الطريقة :-

١- هذه المعايير تفترض أن معدل النمو منتظم طوال السنة الدراسية ، وتفترض أن فترة الثلاث شهور التي تمثل الإجازة الصيفية تمثل نمو شهر واحد من شهور السنة ، وهذا لا يتفق مع حقائق النمو التي توضح أن النمو في القدرة علي القراءة مثلاً ترتبط بالنمو العقلي للتلميذ وهذا النمو مستمر ومتصل طوال العام .

٢- إنه من الصعوبة عمل معايير تمتد إلي مدي كبير من التلاميذ الذين يؤدون الامتحان ، بل نحصل علي معايير الدرجات العليا والدرجات الدنيا بطريق غير مباشر وهو استكمال المنحني في كل من طرفيه الأعلى والأدنى.

٣- إذا فرضنا استكمال المنحني فإنه من الصعب تفسير الدرجات الواقعة في الجزء المستكمل لأنها لا تمثل الواقع وإنما تمثل متوسطاً فرضياً (محمود عبد الحليم ، ١٩٨٠ ، ١١٥-١١٦) .

opeikandl.com