جامعة حماه

كلية التربية

دبلوم تأهيل تربوي

مقرر القياس والتقويم

المحاضرة الخامسة

(تصنيف أدوات القياس والتقويم وشروط صلاحيتها)

د. أسماء الحسن

العام الدراسي: 2020/2019م

تصنيف أدوات القياس والتقويم

تتنوع أدوات القياس والتقويم في ميدان التربية وعلم النفس، وإذا أردنا تصنيف هذه الأدوات والمقاييس ووضعها في فئات لتعرف نقاط التشابه والاختلاف بينها، لوجدنا أنفسنا أمام تصنيفات عدة، وهذا يحدده الأساس الذي ننطلق منه في التصنيف.

أولاً–تصنيف أدوات القياس والتقويم انطلاقاً من الموضوع الذي نقيسه:

يستخدم المقياس لقياس سمة أو صفة معينة تشكل موضوعاً للقياس والتقويم، ويمكن تصنيف المقاييس من حيث الموضوع الذي تقيسه إلى قسمين رئيسين هما:

1–**المقاييس التربوية:** والتي تتصدى لقياس الموضوعات أو الجوانب المرتبطة بالعملية التربوية كاختبارات التحصيل المعيرة وغير المعيرة، وأدوات تقويم المعلم والكتاب والمنهاج.

2–**المقاييس النفسية:** والتي تتصدى لقياس الموضوعات أو الجوانب النفسية عند الفرد كاختبارات الذكاء، ومقاييس الشخصية، ومقاييس الميول والاتجاهات والقيم.

ثانياً–تصنيف أدوات القياس والتقويم انطلاقاً من شروط الإجراء:

يمكن تصنيف المقاييس من حيث شروط الإجراء إلى قسمين رئيسين هما:

1–المقاييس الفردية:

2–المقاييس الجمعية:

ويبين الجدول الآتي الفروق بين المقاييس الفردية والجمعية:

	المقاييس الفردية	المقاييس الجمعية
عدد الأفراد	تتصدى لقياس سمة أو موضوع معين عند **فرد واحد**	تطبق على **مجموعة من الأفراد** في **آن واحد**
الوقت	تطبيقها يحتاج إلى **وقت طويل**	تطبيقها **لا** يحتاج إلى وقت طويل
الأدوات	يتطلب المقياس غرفة وأدوات خاصة وتوفير شروط ملائمة	**لا** يتطلب المقياس غرفة وأدوات خاصة وتوفير شروط ملائمة
التفاعل	تسمح **بالتفاعل** بين الفاحص والمفحوص	**لا** تسمح بالتفاعل بين الفاحص والمفحوص
ملاحظة التغيرات	تتيح **ملاحظة التغيرات** التي تطرأ على المفحوص أثناء الاختبار.	**لا** تتيح ملاحظة التغيرات التي تطرأ على المفحوص أثناء الاختبار
أمثلة	بعض اختبارات الذكاء، وبعض اختبارات القدرات.	اختبارات التحصيل المعيرة، اختبارات الذكاء الجماعية، اختبارات الميول والاتجاهات والقيم.

ثالثاً–تصنيف أدوات القياس والتقويم انطلاقاً من حيث كيفية بنائها:

يمكن تصنيف المقاييس من حيث كيفية بنائها إلى قسمين رئيسين هما:

1–**المقاييس غير المعيرة:** هي **المقاييس التي يعدها الفاحص لقياس سمة معينة**، ويتعرف إلى مدى تقدم المفحوصين ومعرفة نقاط القوة والضعف لديهم، ومدى تقدمهم لتحقيق الأهداف، ومن أمثلتها الاختبارات التي يعدها المعلم.

2–**المقاييس المعيرة: هي الاختبارات التي يعدها مجموعة من الخبراء والمختصين، يتم فيها توحيد شروط التطبيق والتصحيح** على عينة التقنين، و**يتم التوصل إلى معايير**، ويتم التدريب على استخدام هذه الاختبارات، ومن أمثلتها اختبارات الذكاء، اختبارات الشخصية، اختبارات التحصيل المعيرة.

رابعاً–تصنيف أدوات القياس والتقويم انطلاقاً من شكل الاستجابة أو المهمة:

يمكن تصنيف المقاييس انطلاقاً من شكل الاستجابة إلى قسمين رئيسين هما:

1–**مقاييس لفظية:** تتطلب استخدام اللغة (المنطوقة أو المكتوبة) سواء في تعليمات المقياس أو مفرداته، مثل اختبارات التحصيل والقدرات، مقاييس الشخصية.

2–**مقاييس أدائية:** والتي تضم أشكال ورسوم أو صور أو متاهات أو مواد عيانية، ولا تتطلب استخدام اللغة (المكتوبة أو الشفوية) أي لا تتطلب إجابات كتابية أو شفهية، وإنما تتطلب أداء معيناً (نشاطاً) من قبل المفحوص، مثل بعض اختبارات الذكاء الفردية، والاستعدادات الميكانيكية. وتناسب هذه المقاييس الأطفال الصغار والأميين والأجانب.

ولابد من الإشارة أن هناك بعض المقاييس التي تنطوي على أجزاء لفظية وأجزاء غير لفظية مثل بعض اختبارات الذكاء (اختبار وكسلر، اختبار ستانفورد–بينيه).

خامساً–تصنيف أدوات القياس والتقويم انطلاقاً من درجة التحديد في المهمات:

يعرف المقياس بأنه عينة من المثيرات أو المهمات تتطلب عينة من الاستجابات، وانطلاقاً من ذلك يمكن تصنيف المقاييس إلى ثلاثة أشكال هي:

1–**مقاييس محددة البناء:** تتألف من عدد من الأسئلة أو العبارات، وعلى المفحوص الإجابة عن كل سؤال أو عبارة ضمن احتمالات محددة مثل (صح، خطأ) أو اختبار إجابة من إجابات عدة (أ، ب، ج، د)، مثل اختبار مينيسوتا واختبارات التحصيل الموضوعية، مقاييس الميول والاتجاهات.

2–**مقاييس محددة البناء جزئياً:** يكون فيها المثير (السؤال) واضحاً ومحدداً، وتترك الحرية للمفحوص

في تأليف الإجابة وتحديد شكلها، مثل الاختبارات المقالية، اختبار تفهم الموضوع.

3-مقاييس غير محددة البناء: تسمى الاختبارات الإضفائية، ويكون المثير أو السؤال غامض ويترك للمفحوص حرية فهم المثير الذي ليس له دلالة محددة، والإجابة عنه غير محددة (غير محدد البناء)، مثل اختبار رورشاخ لبقع الحبر.

سادساً-تصنيف أدوات القياس والتقويم انطلاقاً من القدرة على الإنجاز:

يمكن تصنيف المقاييس انطلاقاً من القدرة على الإنجاز إلى قسمين رئيسين هما:

1-مقاييس السرعة: تهدف إلى رصد ما يستطيع الفرد إنجازه من مهمات خلال فترة محددة، أي تحديد السرعة القصوى للمفحوص في أدائه الاختباري، فزمن الإجابة يكون محدد تحديداً دقيقاً ولا يسمح للفرد بتخطي الزمن، وتتميز الأسئلة بسهولتها.

2-مقاييس الطاقة أو القوة أو التحمل: والتي تهدف إلى تحديد ما يستطيع الفرد إنجازه من مهام دون النظر إلى الوقت، أي تحديد مستوى القدرة أو التحصيل للفرد، وتتميز الأسئلة بصعوبتها.

سابعاً-تصنيف أدوات القياس والتقويم انطلاقاً من الحكم ووضع الدرجة:

يمكن تصنيف المقاييس انطلاقاً من الحكم ووضع الدرجة إلى قسمين رئيسين هما:

1-مقاييس معيارية المرجع: والتي تقارن أداء الفرد مع متوسط أداء المجموعة، ويتغير ترتيب الفرد بتغير المجموعة التي ينتمي إليها، أو بدخول أفراد إليها أو خروجهم.

2-مقاييس محكية المرجع: والتي تقارن أداء الفرد مع إجابة مثالية موضوعة بشكل مسبق (محك الإتقان) دون النظر إلى أداء المجموعة.

ثامناً-تصنيف أدوات القياس والتقويم انطلاقاً من الهدف أو الغرض:

تصنف المقاييس انطلاقاً من الهدف أو الغرض إلى ثلاثة أقسام هي:

1-الاختبارات التمهيدية (القبلية): والتي تتم **قبل** تنفيذ البرنامج أو النشاط، والهدف منها جمع معلومات عن المشاركين ومستواهم واتجاهاتهم ومدى ملائمة البرنامج لهم

2-الاختبارات البنائية (التكوينية أو المرحلية): والتي تتم **أثناء** تنفيذ البرنامج أو النشاط، والهدف منها معرفة مدى تقدم المتعلمين وتقديم معلومات مستمرة لتعديل البرنامج وتحسينه

3-الاختبارات النهائية (البعدية): والتي تتم **بعد** الانتهاء من البرنامج أو النشاط، والهدف منها معرفة الفاعلية الكلية للبرنامج، ومدى تحقق أهدافه، وتقويم البرنامج ككل من حيث تعديله أو استمراره.

شروط صلاحية أدوات القياس والتقويم

تعد أداة القياس أحد عناصر عملية القياس والتي لابد من التحقق من شروط صلاحيتها، قبل استخدامها ولا سيما عند اتخاذ قرارات مهمة ترتبط بالأفراد، ويمكن إجمالي شروط صلاحية أدوات القياس والتقويم في ثلاث شروط هي: **الصدق، الثبات، القابلية للاستعمال.**

الصدق

معنى الصدق:

يشير الصدق إلى قدرة المقياس على قياس ما وضع لقياسه أو ما نريد نحن قياسه، ويعرف بأنه درجة قياس المقياس لما وضع لقياسه.

فاختبار الذكاء يجب أن يقيس الذكاء وليس شيء أخر، وبقدر ما يقيس هذه السمة بقدر ما يكون صادقاً، والاختبار التحصيلي لمادة الرياضيات للصف الرابع يجب أن يقيس هذه السمة التي وضع لقياسها.

خصائص الصدق:

يتصف الصدق بمجموعة من الخصائص هي:

1-يرتبط الصدق ببنية المقياس ونتائجه.

2-يرتبط الصدق بأغراض الاختبار واستخدامه: فالاختبار الذي يتصف بدرجة مرتفعة من الصدق لغرض معين، قد لا يكون كذلك بالنسبة لغرض آخر، فاختبار القدرة القرائية لا يصلح لقياس القدرة الحسابية والاختبار المعد للتشخيص وبدرجة عالية من الصدق قد لا يصلح للتنبؤ بالأداء، ومنه تضعف درجة صدقه.

3-يرتبط صدق المقياس بالعينة التي وضع لها: فالاختبار المعد لقياس سمة معينة عند مجموعة أو فئة معينة-حسب العمر أو الجنس أو...الخ-وعلى درجة عالية من الصدق قد ينخفض صدقه عند تطبيقه على مجموعة أو فئة أخرى.

4-الصدق على درجات وهو نسبي: فلا نستطيع القول إن الاختبار أو المقياس صادق تماماً أو غير صادق أبداً، وإنما يمكننا الحديث عن **درجة وصفية** مثل مرتفع، متوسط، منخفض، وقد تستخدم الأرقام لتقدم **درجة كمية** من مقدار الصدق.

أشكال الصدق:

تتعدد أشكال الصدق ويمكن التمييز بين ثلاثة أنواع رئيسة للصدق وهي (صدق المحتوى، الصدق المحكي، الصدق البنيوي) إضافة إلى شكل رابع وهو (الصدق الظاهري).

أولاً–صدق المحتوى:

يعرف بأنه: مدى تمثيل محتوى الاختبار أو المقياس للسمة أو الموضوع المراد الاستدلال عليه أو قياسه. ويسمى "**بالصدق التمثيلي**" لأنه يعتمد على مدى تمثيل الاختبار للسمة المقيسة أو إلى أي مدى يعد الاختبار عينة ممثلة لتلك السمة، ويسمى أيضاً "**بالصدق المنطقي**" لأنه يعتمد على التحليل المنطقي لبنود المقياس ومحتواه، ولا يعتمد على التحليل الكمي. ويعد هذا الصدق شرط ضروري لكل المقاييس ولكنه غير كافٍ. ومن طرائق حساب صدق المحتوى عرضه على مجموعة من المحكمين المتخصصين في مجال موضوع المقياس وفي القياس والتقويم.

ويعد هذا الصدق مهماً لكل أنواع المقاييس والاختبارات ولا سيما الاختبارات التحصيلية، ويتم التحقق من صدق المحتوى للاختبارات التحصيلية من خلال عرضه على مجموعة من المحكمين إضافة إلى جدول المواصفات.

ثانياً–الصدق المحكي:

يعتمد الصدق المحكي على دراسة الارتباط أو العلاقة بين درجات الاختبار أو المقياس ودرجات مقياس آخر يعتمد كمحك. ويعرف الصدق المحكي بأنه: **الدرجة التي يترابط عندها الأداء على مقياس مع الأداء على مقياس آخر اعتمد محكاً للمقياس الأول وأساساً في الحكم على صدقه.**

ويسمى الصدق المحكي **بالصدق التجريبي أو الصدق الإحصائي** لأنه يعتمد على التجريب والإحصاء، ومن خلال هذا الصدق نحصل على درجة كمية تدلنا على مقدار هذا الشكل من الصدق.

ومنه فإن الصدق المحكي يدرس العلاقة بين درجات اختبارات ومقاييس تعد بمثابة منبئات ودرجات مقاييس خارجية مستقلة تعد بمثابة محكات، وتتراوح قيمة الصدق المحكي بين –1 و +1.

ومن الصعوبات في الصدق المحكي إيجاد المحكات الملائمة، لأن دقة تقدير الصدق المحكي تعتمد على نوع المحك المستخدم.

أنواع الصدق المحكي:

وله نوعان: الصدق التنبؤي والصدق التلازمي (المصاحب)، ويبين الجدول الآتي أوجه الشبه والاختلاف بين هذين النوعين:

الصدق التنبؤي	الصدق التلازمي (المصاحب)
كلاهما نوعان من الصدق المحكي، ويتطلبان وجود محك خارجي	
كلاهما يعتمدان على **التحليل الكمي** وليس الكيفي أو المنطقي	
كلاهما يستخدم **معامل الارتباط** بين درجات الأفراد على المقياس وعلى المحك، ودرجة الترابط تدل على مقدار الصدق	
يركز على دراسة العلاقة بين درجات الأفراد على الاختبار ودرجاتهم على اختبار آخر (محك) يطبق **بعد فترة زمنية** من تطبيق الاختبار الأول	يركز على دراسة العلاقة بين درجات الأفراد على الاختبار ودرجاتهم على اختبار آخر (محك) بحيث **لا يوجد فاصل زمني (أو فاصل زمني قصير)** بين تطبيق الاختبارين
يهدف إلى التنبؤ بأداء المفحوصين وإنجازاتهم في المستقبل، وتعرف القدرة التنبؤية للمقياس، إذ يعد مهماً جداً وخاصة عند اتخاذ قرارات مهمة مرتبطة بالأفراد	يهدف إلى التعرف إلى الأداء الراهن للفرد وتشخيصه، وتعرف دقة المقياس في القياس السمة موضع القياس
يلائم بصورة خاصة اختبارات القدرات والاستعدادات	يلائم بصورة خاصة اختبارات الشخصية
قيم الصدق التنبؤي ليست مرتفعة نتيجة تعقد السلوك الإنساني وتشابكه	قيم الصدق التلازمي أعلى من قيم الصدق التنبؤي

خطوات الصدق التنبؤي:

1-تطبيق المقياس الذي نرغب بمعرفة درجة صدقه على عينة من الأفراد ويسمى المقياس الأولي أو المتنبَئ به.

2-حساب درجات الأفراد على المقياس الأولي

3-إخضاع أفراد العينة لبرنامج تدريبي أو تعليمي لفترة زمنية

4-تطبيق اختبار أو مقياس ثانٍ على أفراد العينة

5-حساب درجات الأفراد على المقياس الثاني (المحك)

6-حساب الارتباط بين أداء أفراد العينة على الاختبارين أو المقياسين

إن الصدق التلازمي كالصدق التنبؤي في خطواته ولكن لا يوجد فترة زمنية تفصل بين الاختبارين الأول والثاني.

ثالثاً-الصدق البنيوي أو التكوين الفرضي (الافتراضي):

يعرف بأنه: الدرجة التي يقيس بها الاختبار السمة أو الخاصية التي يفترض أن يقيسها، ويهتم بتعرف مدى اتفاق درجات الاختبار مع نظرية معينة أو مجموعة من المكونات المنبثقة عن نظرية في المجال. ويعد مفهوم شامل لكافة أنواع الصدق الأخرى وشرط ضروري للاختبارات النفسية والتربوية.

ويعتمد على التحليل المنطقي والوسائل التجريبية والإحصائية، فبداية لا بد من إعداد بنود تقيس مكونات أو أجزاء السمة المقاسة ويعد هذا صدقاً للمحتوى، ثم يتم تجريب الاختبار على عينة ممثلة للمجتمع وإجراء تحليل عاملي للتحقق من قياسه للبنية المفترضة، وإجراء دراسات ارتباطية بين درجات الاختبار ودرجات محكات خارجية لتعرف مدى جودة درجات الاختبار في تمثيل مكوناته وقياسها.

رابعاً-الصدق الظاهري:

لا يشير إلى ما يقيسه الاختبار فعلاً، بل إلى ما يبدو **ظاهرياً** أنه يقيسه سواء في نظر المفحوصين أنفسهم أم في نظر غيرهم من غير المختصين. أي يشير إلى مدى تقبل المفحوصين وغيرهم من غير المختصين للاختبار أو رفضهم له. ومنه فإنه يرتبط **بنظرة الأفراد تجاه المقياس** من حيث دافعيتهم لاستخدامه أو حكمهم عليه من خلال شكله وقناعتهم بنتائجه وقدرته على التنبؤ ودوره في اتخاذ القرارات. لذلك لا بد من استثمار دافعية المفحوصين لتأمين الصدق الظاهري وضمان تعاونهم في الموقف الاختباري، ويتم ذلك بإعادة صياغة بعض العبارات، توضيح التعليمات، وطريقة استخدام الاختبار، وزيادة التشويق بما يتناسب وأعمار المفحوصين ليظهر الاختبار أكثر ارتباطاً بموقف القياس الفعلي.

العوامل المؤثرة في صدق الاختبار:

هناك عوامل كثيرة تؤثر في صدق الاختبار وهذه العوامل كما صنفها جرونلند هي:

1-عوامل تتصل بالاختبار نفسه:

أ-المفردات والتراكيب الصعبة: والتي تؤدي إلى عدم قدرة المفحوصين معرفة المطلوب بدقة من هذه البنود، وبالتالي لا يقيس الاختبار ما وضع لقياسه.

ب-البنود الصعبة جداً والسهلة جداً: لأنها تؤدي إلى ضعف القدرة التمييزية للاختبار.

ج-عدد البنود: كلما زاد عدد البنود زاد تمثيل السمة المقاسة وبالتالي ارتفع الصدق، والعكس صحيح.

د-البنود الموحية بالإجابة: تؤدي إلى دفع المفحوص نحو إجابة معينة وبالتالي تضعف درجة الصدق.

هـ-التعليمات غير الواضحة: والتي تضلل المفحوص لإعطاء الاستجابة بالطريقة المناسبة.

2-عوامل تتصل بشروط الإجراء والتصحيح:

أ-الشروط البيئية: كالطقس والحرارة والبرودة والرطوبة والإضاءة والصوت، فإذا كانت غير مناسبة فإنها تؤثر سلباً على أداء المفحوصين وتضعف الصدق.

ب-الوقت المخصص للاختبار: كلما زاد وقت الاختبار زاد الصدق إلى حد معين ثم يبدأ بالتناقص. فالوقت الضيق يسبب ارباك المفحوص، والوقت الطويل يؤدي إلى الملل والغش.

ج–الغش: يضعف الصدق لأنه يسمح بالتقدير الخاطئ للسمة المقاسة.

د–أخطاء التصحيح: فنحصل على تقديرات لا تمثل السمة أو الصفة لدى المفحوصين.

3–عوامل تتصل باستجابات المفحوصين: كالأزمات الانفعالية والإرهاق والتعب وضعف الدافعية.

4–عوامل تتصل بطبيعة المجموعة والمحك: فالاختبار الذي يكون على درجة عالية من الصدق مع جماعة قد لا يكون كذلك مع جماعة أخرى، وبالنسبة للمحك كلما زاد التشابه بين السلوكيات التي يقيسها الاختبار والسلوكيات التي يقيسها المحك كلما ارتفع الصدق المحكي.

الثبات "الموثوقية"

معنى الثبات:

يعرف الثبات بأنه: درجة ثبات **النتائج** في حال تكرار القياس أكثر من مرة وعلى نفس أفراد العينة. فالثبات يرتبط بالنتائج فقط ولا يرتبط بالمقياس أبداً.

خصائص الثبات:

1–الثبات مفهوم إحصائي كمي محض، بخلاف الصق الذي يعتمد على التحليل الكمي والكيفي.

2–يرتبط الثبات بالنتائج وليس بالأداة بخلاف الصدق الذي يرتبط بالنتائج وبالأداة.

3–ترتبط طريقة تقدير الثبات بالغرض أو بالهدف من استخدام الأداة، بخلاف الصدق والذي يمكن فيه استخدام الطرائق كافة مهما كان الغرض من الأداة.

العلاقة بين الصدق والثبات:

كلا الصدق والثبات شرطان ضروريان من شروط صلاحية المقياس، وإن الثبات شرط ضروري ولكنه غير كافٍ. وإن كل اختبار يتصف بالصدق يتصف بالثبات، ولكن ليس بالضرورة أن كل اختبار ثابت هو اختبار صادق، ومنه كلما ارتفعت درجة صدق الاختبار ارتفعت درجة ثبات نتائجه وليس بالضرورة إذا ارتفعت درجة ثبات نتائجه ارتفعت درجة صدقه.

طرائق حساب الثبات:

هناك طرائق عدة لحساب الثبات منها:

1–طريقة الإعادة:

يتم في هذه الطريقة تطبيق الاختبار على عينة من الأفراد ثم تطبيقه مرة ثانية عليهم بعد فترة زمنية، قد تصل إلى أسابيع، ثم حساب معامل الارتباط بين درجات أفراد العينة على الاختبار في المرتين، وتمثل

هذه الطريقة معنى الثبات تماماً إذ أننا نستخدم نفس الاختبار ونفس المجموعة في مرتي التطبيق، ويسمى معامل الثبات الناتج "**بمعامل الاستقرار**".

ومن عيوب هذه الطريقة:

- التأثير الناتج عن الفترة الزمنية للأفراد، من خلال اكتساب معارف جديدة أو النسيان.
- تعد هذه الطريقة غير اقتصادية في الوقت والجهد.
- قد تثير هذه الطريقة استغراب أو استهجان أو ملل أفراد المجموعة أثناء التطبيق مرة ثانية.

2-طريقة الأشكال المتعادلة أو المتكافئة:

يتم في هذه الطريقة تطبيق الاختبار على مجموعة من الأفراد، وإيجاد اختبار ثانٍ مكافئ للاختبار الأول من حيث الموضوع وعدد البنود والسهولة والصعوبة والمهمات المطلوبة، ثم حساب معامل الارتباط بين درجات أفراد المجموعة على الاختبارين، ويسمى معامل الثبات الناتج "**بمعامل التعادل أو التكافؤ**".

ومن ميزات هذه الطريقة: عدم الحاجة لتطبيق نفس الاختبار مرة ثانية، وبالتالي استبعاد أثر الفترة الزمنية الفاصلة بين التطبيقين.

ومن عيوب هذه الطريقة:

- صعوبة إيجاد مقياس مكافئ تماماً للمقياس الأول.
- الوقت والجهد الكبيرين في إيجاد المقياس المكافئ.
- ما نحصل عليه ليس معامل ثبات وإنما معامل التكافؤ بين الاختبارين.

3-طريقة التنصيف أو التجزئة النصفية:

يتم في هذه الطريقة تطبيق الاختبار على مجموعة من الأفراد، ثم تقسيم بنود الاختبار إلى قسمين متعادلين ومتكافئين، ويتم تقسيم البنود وفق طريقتين: إما حسب البنود الفردية والزوجية، أو حسب معاملات السهولة والصعوبة للبنود. ثم حساب معامل الارتباط بين درجات الأفراد على القسمين، ويسمى معامل الثبات الناتج "**بمعامل الاتساق أو التجانس الداخلي**".

ومن ميزات هذه الطريقة أنها:

- تختصر الوقت والجهد الكلفة.
- عدم تأثر المفحوصين بالفترة الزمنية إذ يطبق الاختبار مرة واحدة.
- عدم الحاجة إلى اختبارين متكافئين.

ومن عيوب هذه الطريقة:

- يبين معامل الثبات المحسوب مدى ارتباط كل قسم بالآخر، وليس ثبات نتائج الاختبار ككل.
- صعوبة تقسيم الاختبار إلى قسمين متعادلين ومتكافئين تماماً، من حيث المهمات والصعوبة.
- إن تقسيم الاختبار يضعف تمثيل البنود للسمة المقاسة، وهذا بدوره يضعف صدق الاختبار وبالتالي تضعف درجة ثبات النتائج. لذلك لابد من تصحيح معامل الثبات الناتج باستخدام صيغة إحصائية تعرف باسم معادلة سبيرمان–بروان.

العوامل المؤثرة في ثبات نتائج الاختبار:

من العوامل المؤثرة في ثبات نتائج الاختبار:

1–**عدد البنود:** كلما زاد عدد البنود ارتفعت درجة الصدق ومنه ترتفع درجة الثبات، والعكس صحيح.

2–**الزمن المخصص للاختبار:** كلما زاد زمن الاختبار ارتفعت درجة الثبات، ولكن إلى حد معين ثم يبدأ بالتناقص، نتيجة الملل أو الغش عند المفحوصين.

3–**التباين بين الدرجات:** كلما زاد التباين بين الدرجات أدى إلى عدم تذبذب مراكز الأفراد داخل المجموعة، وهذا يؤدي إلى ارتفاع درجة الثبات.

4–**الحزر أو التخمين:** فالتخمين يؤدي إلى ضعف الثبات، لأن التخمينات المفحوصين تختلف بين المرة الأولى والثانية.

5–**صيغة البنود:** فالبنود الغامضة أو الطويلة تقلل من درجة الثبات، على عكس البنود الواضحة.

6–**حالة الفرد النفسية والصحية:** فالتعب والإرهاق والتوتر والاضطراب والانزعاج يؤدي إلى عدم قدرة الفرد على أداء موازٍ في الاختبار الأول والثاني.

القابلية للاستعمال "الكفاية"

تعد القابلية للاستعمال "**الكفاية**" الشرط الثالث من شروط صلاحية المقياس، وترتبط بثلاثة جوانب رئيسة خارجة عن نطاق الاختبار هي: **الوقت والجهد النفقات.** ولها أهمية كبيرة، وتعد شرطاً مهماً في الاختبارات التي تطبق على أعداد كبيرة من الأفراد حيث يكون هناك حساب للوقت والجهد والنفقات.

ولكن القابلية للاستعمال يجب ألا تعطى الأولوية في صلاحية المقياس إذا ما قورنت بالصدق والثبات، فهما الأهم مهما كانت الكلفة والوقت والجهد.

ومن الجوانب المهمة في القابلية للاستعمال: **وضوح التعليمات، توقيت إجراء الاختبار، وتوفير الشروط البيئية المناسبة،** لأنها تساعد على أداء المفحوصين بشكل جيد وترفع درجة الصدق والثبات.