



وزارة التعليم العالي والبحث العلمي

جامعة تكريت/ كلية التربية للعلوم الإنسانية

قسم العلوم النفسية والتربوية/ علم النفس التربوي

القياس والتقويم

خصائص الاختبارات الجيدة

اعداد

١. د صباح مرشود منوخ

مواصفات الاختبار الجيد

لا يعد الاختبار اداة صالحة للقياس الا اذا توافرت فيه شروط معينة وتعد هذه الشروط بمثابة اهداف يحاول مصمم الاختبار تحقيقها اثناء تصميمه الاختبار، واهم هذه الشروط صدق الاختبار، ويليها ثبات الاختبار ومن ثم موضوعيته وكذلك توفر بعض العوامل التي تجعل الاختبار ملائما للتطبيق العملي كشموليته وسهولة استعماله.

أولاً. الصدق:

يعد الصدق من الأمور التي يجب ان يتأكد منها مصمم الاختبار عندما يريد بناء اختبار، فلا بد من ان تكون لديه ظاهرة سلوكية محددة لقياسها، كان يقيس القيم الاجتماعية او سمات الشخصية او الميول المهنية او الانبساط او الانطواء او الذكاء او التعصب او التحصيل الدراسي في مادة معينة، فيقوم بترجمة الظاهرة السلوكية الى عبارات او مواقف يعتمد كفقرات للاختبار يسمى باختبار سمات الشخصية او القيم او ما شابه ذلك، وبعد التأكد بطرق علمية من ان الاختبار يقيس ما وضع من اجله او السمة المراد قياسها، فالاختبار الصادق اذن هو ذلك الاختبار القادر على قياس السمة التي وضع من اجلها فمثلا يعد اختبار الاستعداد المدرسي صادقا اذا كان قادرا على قياس الاستعداد الدراسي، ويكون غير صادق اذا كان يقيس ظاهرة سلوكية أخرى، وبذلك يقال ان الاختبار صادق اذا توفرت الأدلة الكافية على انه يقيس فعلا العامل الذي صمم الاختبار لقياسه، وانه لا يقيس ايه عوامل أخرى.

أنواع الصدق:

١. الصدق الظاهري:

وهو اقل أنواع الصدق دقة، ويستخدم عادة في الاختبارات غير المقننة وإذا لم يسبق ان اختبرت درجة صدقها فإنه يؤول الى ان يكون ضعيفا، وإذا كان وجود هذا النوع من الصدق واجبا فأننا يجب ان ننظر اليه بتحفظ، والصدق الظاهري هو المظهر العام للاختبار، أي الاطار الخارجي له ويشمل نوع المفردات وكيفية صياغتها ووضوحها ودرجة موضوعيتها.

٢. صدق المضمون:

ويعد من الخطوات الأساسية في تصميم الاختبار، لان تصميم أي اختبار يقوم على تحديد السمة او الظاهرة المراد قياسها منطقيا، ثم اجراء تحليل لذلك لشموليته بحيث تتضح اقسامه ويتضح ترتيبها بحسب أهميتها، ثم تحديد الاوزان بكل قسم منها تمهيدا لوضع مفردات الاختبار بحيث تكون ملائمة لهذه الاوزان، وصدق المضمون يهتم بمفردات الاختبار ومحتوياته ومادته.

٣. الصدق الداخلي:

يعتمد هذا النوع من الصدق على تحليل محتوى الاختبار بطرائق إحصائية لمعرفة درجة السهولة والصعوبة ومدى قدرة هذه العناصر على التمييز بين الافراد الذين يحصلون على درجات عالية في الاختبار، والذين يحصلون على درجات منخفضة في الاختبار نفسه، وبهذه الطريقة يمكن تحسين عناصر الاختبار واختيار العناصر التي لها القدرة التمييزية بين الافراد.

٤. الصدق التنبؤي:

ويتم عن طريق استخدام نتائج اختبار معين للتنبؤ بسلوك معين، وتتخذ هذه الصفة التنبؤية للاختبار دليلا على صدقه والاختبار القادر على التشخيص او التنبؤ بالسلوك المقصود يقال عنه ان له صدقا تنبؤيا، وبمعنى اخر هناك علاقة بين نتائج الاختبار في الوقت الحاضر والصورة المستقبلية، ومن امثلة هذا النوع اختبارات الذكاء فهي ضرورية للتنبؤ بالنجاح الأكاديمي.

٥. الصدق التلازمي:

ويمكن إيجاد هذا النوع من الصدق بمقارنة النتائج التي يحصل عليها الباحث من الاختبار بالنتائج التي يحصل عليها من مقاييس أخرى لنفس الأداء، تطبق في الوقت نفسه، ومنها تقديرات المدرسين او أساليب أخرى تم اثبات صدقها قبلا.

٦. صدق البناء:

اكثر أنواع الصدق تعقيدا، لا يهتم هذا النوع بأسلوب القياس فقط، وانما يهتم بالنظرية المطروحة في ضوءها وتفسير النتائج التي يحصل عليها من استخدامها، وعلى الباحث الذي يسعى لتحقيق من صدق تكوين الاختبار فحصه، ثم اقتراح تكوينات قادرة على إعطاء تفسيرات لأنواع الأداء على الاختبار، ومن ثم اشتقاق الفروض من النظرية التي صمم في ضوءها، ويتم اختبار الفروض من خلال اجراء دراسات تجريبية عنها.

٧. الصدق العاملي:

ويعتمد هذا النوع من الصدق على التحليل العاملي للاختبارات المختلفة ولموازينها وتقوم الفكرة على حساب معاملات ارتباط الاختبارات والموازن

المختلفة، ثم تحليل هذه الارتباطات الى العوامل التي أدت الى ظهورها مما يؤدي الى الكشف عن العوامل العامة المشتركة والطائفية التي تتكون منها الاختبارات المختلفة اذ ان العامل العام يؤثر في جميع الاختبارات وبنسب مختلفة، بينما العامل الطائفي يؤثر في بعض الاختبارات بنسب مختلفة أيضا وتسمى معاملات تشبع الاختبارات بالعامل

ثانيا. الثبات

ونعني به التوصل الى النتائج نفسها عند تطبيق الاختبار في مدتين مختلفتين وفي حدود زمن يتراوح بين أسبوع واسبوعين عادة اذ ان قلة المدة عن ذلك يتيح فرصة للتذكر، كما ان طول المدة قد يتيح فرصة لنمو الافراد ومن ثم تغير أدائهم.

يتم حساب معامل الثبات عمليا عن طريق تطبيق الاختبار في فرصتين مختلفتين وهناك طرائق مختلفة لحساب معامل الثبات، ولكن كلها تستند الى الفروق الفردية التي تتكون من الاختبار يرجع جزء منها الى أخطاء القياس والجزء الاخر الى فروق فردية حقيقية.

وعموما ليس هناك معامل مطلق للثبات، اذ يتوقف على نوع المؤثرات التي يمكن اعتبارها اخطاءا تجريبية في البحوث ويتوقف ذلك على طبيعة البحث وفروضه وأهدافه وهناك عدة طرائق لحساب معامل ثبات الاختبار هي:

١. إعادة الاختبار : Test –retest

يسمى معامل الثبات الناتج من هذه الطريقة بمعامل الاستقرار، أي استقرار نتائج الاختبار خلال الفترة بين التطبيق الأول والثاني للاختبار، حيث ان هذا النوع في حقيقته هو معامل الارتباط بين مجموعتي درجات الاختبار في

التطبيقات، وذلك بأعطاء نفس مجموعة الافراد الاختبار نفسه مرتين، ومصدر الخطأ في الدرجة التي يحصل عليها الفرد في ضوء معامل الثبات هو خطأ التغيرات عبر الفقرة الزمنية.

٢. أسلوب الصور المتكافئة:

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 صورتين الاختبار، فإن الارتباط بين الصورتين سيكون معاملًا للثبات، ويسمى معامل الثبات بهذا الأسلوب بمعامل التكافؤ حيث يدل على مدى الارتباط بين درجات اختبار واختبار آخر متكافئ معه يعطي في الوقت نفسه أو من وقت متقارب جدًا 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.

٣. أسلوب التجزئة النصفية:

يعتمد هذا الأسلوب أساساً على تقسيم فقرات المقياس أو الاختبار إلى 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 فإن أنه يقيس التجانس بين الفقرات، أي الاتساق الداخلي، مما يسمى بمعامل الثبات بهذا الأسلوب (معامل الاتساق الداخلي) وإن تباين الخطأ الذي يقيسه أسلوب التجزئة النصفية هو عدم تجانس نصفي الاختبار.

ومن الواضح ان حساب الثبات بالتجزئة النصفية لا يقيس التجانس الكلي للمقياس لأنه يقسم الفقرات الى قسمين لذلك فهو معامل ثبات لنصف الاختبار او المقياس، ولتلافي ذلك ينبغي تصحيح معامل الثبات المستخرج بطريقة حساب الارتباط، ومنها معادلة (سبيرمان - براون) كما هناك بعض الطرق التي يحسب بها كل معامل الثبات.

٤. تحليل التباين:

ان طريقة التجزئة النصفية تصلح لحساب معامل الاتساق الداخلي (تجانس نصفي الاختبار) لكنها لا تقيس تجانس الأسئلة ذاتها في الاختبار، الا ان بعض المختصين يرون ان تجانس الأسئلة لا يدخل في مجال موضوع (ثبات الاختبارات) وانما هي خاصية من خصائص الاختبارات الجيدة، ومع ذلك لا يمكن استبعاده من مناقشة الثبات، حيث انه يتناول استقرار استجابات المفحوص على الأسئلة او فقرات الاختبار واحدا بعد الاخر، وعادة ما يسمى هذا النوع من الثبات بـ (معامل الاتساق داخل أسئلة او فقرات الاختبار) ويختلف (معامل الاتساق داخل الأسئلة) عن (معامل الاتساق الداخلي) حيث ان معامل الاتساق داخل الأسئلة يشير الى التجانس الكلي للاختبار بالمعنى المباشر، في حين يكون معامل الاتساق الداخلي للاختبار اقرب الى معنى التكافؤ كما يتمثل في الصور المتكافئة للاختبار، فقد يكون لدينا اختبارين لهما معاملان للثبات عاليان تم حسابهما بطريقة الصور المتكافئة او التجزئة النصفية، لكنهما قد يختلفان في درجة تجانس الأسئلة التي يتكون منها صورتا الاختبار او نصفاه.

وعموما فإن الفرق بين معامل الاتساق الداخلي - الذي يحسب بالتجزئة النصفية وبين معامل الاتساق داخل الأسئلة يمكن اختباره دليلا على مدى التجانس أو التباين في أسئلة الاختبار .

٥. ثبات المصححين:

من المصادر الرئيسية لتباين الخطأ في درجة الاختبار أو المقياس الذي يعتمد على تقييم المصحح وليس على مفتاح التصحيح، وهذا ما يحدث عادة في الاختبارات الاسقاطية التي تعتمد أساسا على تاويل استجابة المفحوص للمنبه وفقا لمحاكاة معينة، لكن انطباقه لتأويل على شروط المحاكات كثيرا ما يرجع الى ذاتية المصحح رغم التزامه بشروط المحك، مما تظهر الفروق بوضوح عندما يستخدم مصححان او اكثر نفس المحك لتصحيح نفس العينة من الاختبارات، كما تظهر نفس الحالة أيضا في الاختبارات الخاصة بقياس القدرات الإبداعية، حيث ان المصحح هو الذي يحدد مدى جودة او اصالة الاستجابة في اختبارات الاصاله رغم توفر معايير التصحيح كما ان الاختبارات التحصيلية لاسيما في اختبار المقال يمكن تقدير درجتها على الحكم الذاتي للمصحح مما يمكن ان تصحح إجابات المفحوصين بواسطة اثنين او اكثر من الحكام او الخبراء، ومن ثم إيجاد معامل ثبات المصححين على نفس المجموعة، بيد انه نادرا ما تحتاج الاختبارات الموضوعية الى حساب معامل ثبات المصححين، وذلك يعود لقله أخطاء التصحيح فيها ولسهولتها في إعطاء الدرجة وتقدير عدد وحجم الأخطاء .

تفسير معامل الثبات:

لكل معامل ثبات تفسير يختلف باختلاف الطريقة التي استخرج بها فأذا حسبنا معامل الثبات بطريقة إعادة الاختبار فان معامل الثبات يشير الى

مدى استقرار علامات الطلاب، ولكن عند حساب معامل الثبات باستخدام الصور المتكافئة فإن معامل الثبات يشير الى مدى استقرار علامات الطلاب، ولكن عند حساب معامل الثبات باستخدام الصور المتكافئة فإن معامل الثبات يشير الى مدى تكافؤ علامات الطلاب في الاختبارين.

وإذا حسب معامل الثبات النصفي، فإن معامل الثبات يشير الى مدى تكافؤ نصفي الاختبار المكونين للاختبار الكلي.

العوامل التي تؤثر على معامل الثبات:

١. طول الاختبار:

لقد رأينا ان معامل الثبات يتأثر بطول الاختبار او بعدد فقرات الاختبار ولقد وجد ان كلما زاد عدد فقرات الاختبار كلما زاد معامل الثبات لحساب معامل الثبات لاختبار اذا جعلناه أطول او اقصر تستخدم معادلة سبيرمان - براون

٢. تباين قدرات الطلاب:

نعني بتباين قدرات الافراد أي وجود فروق بين قدرات الافراد فإذا كان الافراد متجانسين فيكون التباين قليلا وهذا يؤدي الى اختلاف رتبة الفرد بين اقرانه الذين لهم نفس القدرة مما يؤدي الى نقصان معامل الثبات، لذلك كلما زاد التباين بين الافراد كلما زاد معامل ثبات الاختبار.

٣. معامل صعوبة السؤال.

إذا كانت الأسئلة صعبة جدا فهذا يعني ان معظم الطلبة لا يجيبون إجابة صحيحة فيظهر الطلاب وكأنهم متجانسون، كذلك الحال عندما تكون علاماتهم مرتفعة والتباين بين الطلاب قليل وبالتالي فإن معامل الثبات يكون

منخفضا اذا كانت الأسئلة صعبة جدا او اذا كانت الأسئلة سهلة جدا، وحتى يكون معامل ثبات الاختبار عاليا يفضل ان يكون معامل صعوبة الأسئلة محصورا بين ٠,٢٥ و ٠,٧٥، وافضل معامل صعوبة هو ٠,٥٠.

٤. مستوى قدرة الطلاب.

كلما كان الاختبار مناسباً لمستوى قدرة الطلاب كلما كان معامل ثباته اعلى فلو وضعنا اختبارا مناسباً لطلاب الصف السادس في مادة الرياضيات فأنهم سيجيبون عليه لانه يناسبهم بينما لو طبقناه على طلاب الصف الخامس فأنهم سيجيبون في الغالب باستخدام التخمين وهذا يقلل من ثبات الاختبار لانه عند إعادة الاختبار فانهم سوف يخمنون إجابات غير الإجابات التي أجابوا عليها في الاختبار الأول وذلك لضعف قدرتهم او لعدم فهمهم السؤال.

ثالثا. الموضوعية:

يقصد بالموضوعية التحرر من التحيز او التعصب، وعدم ادخال العوامل الشخصية فيما يصدر الباحث من احكام وعليه ماذا يمكن توقعه اذا طبق باحثان اختبارا موضوعيا مما لا شك فأنهما سوف يحصلان على نفس النتائج اذا كان الاختبار موضوعيا وهذا يعني ان الباحث الموضوعي يجب ان يلاحظ نفس السلوك، وان يسجل ملاحظاته أولا بأول حتى لا يتأثر استرجاعه بالنسيان او الخطأ، وان يقدر السلوك متبعاً نفس القواعد والموضوعية في تقويم الإجابات معناه ان الدرجة لا تتأثر بالاحكام الذاتية للشخص الذي يقوم الإجابات وتوفر الموضوعية يساهم في ثبات درجات الاختبار ودقتها.

والاختبارات التي يطلب من المفحوص اختبار استجابة واحدة من عدد من الاستجابات تسمى بالاختبارات الموضوعية ومن أمثلتها الاختيار المتعدد والمطابقة والخطأ والصواب.

رابعاً. السهولة:

ونعني بسهولة الاختبار أي سهولة اجراء الاختبار وتطبيقه وبذلك من خلال تضمين الاختبار تعليمات وافية وواضحة عن كيفية اخذ الاختبار وان تكون هذه التعليمات محتوية على نماذج توضح نوع الأسئلة وكيفية الإجابة عليها وان تحتوي على تمارين اذا كانت الأسئلة من نوع جديد بالنسبة للمبحوثين.

خامساً. الشمولية:

ونعني بالشمولية ان يحصل الممتحن على اكبر قدر من الشمول في التقدير الذي يجريه لقدرة الطالب، ان درجة الشمول وتوثر عادة في مدى ثباتها وصدقها فإذا وجه سؤال واحد فقط الى احد التلاميذ ويجب عليه إجابة صحيحة فانه يحصل على درجة كاملة وإذا أعيد الاختبار ذاته ووجه الى التلميذ سؤال آخر غيره وعجز عن الإجابة فانه يحصل على صفر، وبما ان مثل هذه المواقف متوقعة الحدوث نتيجة لعوامل الصدفة فان معيار الشمول يلعب دوراً مهماً في الوصول الى نتائج منسجمة وثابتة.

سادساً. التقنين والمعايير:

ومن صفات الاختبار الجيد ان تكون إجراءات تطبيقه على الافراد مقننة، لذا نجد ان الاختبارات المقننة تحتوي على تعليمات لمن يريد تطبيق الاختبار، توضع له كيفية التطبيق بشكل محدد تماماً وذلك لضمان تساوي او تكافؤ

إجراءات التطبيق بتفصيلاتها المحددة، وفي كل مرة يطبق فيها الاختبار
مهما اختلف الأشخاص الذين يشرفون على التطبيق

المصادر:

١. مجيد ، سوسن شاكر (٢٠١٤): أسس بناء الاختبارات والمقاييس النفسية والتربوية، جامعة بغداد ، كلية التربية، مركز ديونو لتعليم التفكير.
٢. العجيلي، صباح حسين (١٩٩٠)، التقويم والقياس، جامعة بغداد، كلية التربية ابن رشد، دار الحكمة، بغداد.
٣. حنا، مناهج البحث التربوي، جامعة بغداد.