المحاضرة السابعة: مواصفات الاختبار الجيد

تمهيد:

تُعتبر الاختبارات المدرسية من أهم الوسائل التي يعتمد عليها في قياس وتقويم قدرات الطلاب؛ وذلك من أجل تحديد مستواهم التحصيلي، ومعرفة مدى تحقيق الأهداف السلوكية أو النواتج التعليمية، وما يقدمه المعلم من نشاطات تعليمية مختلفة تساعد على رفع الكفايات التحصيلية لدى الطلاب؛ لذلك حرص المشرفون التربويون على أن تكون هذه الاختبارات ذات كفاءة عالية في عملية القياس والتقويم، وحتى تحقق ذلك لابد من توفرها على مواصفات محددة تجعل منها وسيلة صالحة للتقويم وهذه المواصفات تتمثل فيما يلي: الموضوعية، الثبات، الصدق، الشمولية، معامل التمييز، معامل الصعوبة والسهولة، القابلية للاستعمال...

1. مواصفات الاختبار الجيد:

يؤكد علماء القياس على أن خاصيتي الثبات والصدق من أهم خصائص أداة القياس الجيدة وسماتها فبدونهما لا يمكن الوثوق في قدرة الأداة على قياس ما صممت لقياسه ولا بدقة النتائج المتحصل عليها عند استخدامها لقياس السمات المختلفة.

وهذا لا يعني إهمال الخصائص والسمات الأخرى التي يجب أن تتمتع بها أداة القياس فهي بالإضافة إلى ذلك يجب أن تكون:

- شاملة وممثلة لجميع مكونات القدرة أو الخاصية المطلوب قياسها؛
- يجب أن تبنى وتحلل بطريقة موضوعية مما يعني عدم تدخل العوامل الذاتية في بنائها وتحليلها وتصحيحها، وبالتالي فإن درجات الفرد ستبقى كما هي حتى باختلاف المصحح؛
- أن تكون مفردة الأداة متصلة بالموضوع المراد قياسه اتصالا جيدا مما يؤدي إلى إيجاد مدى واسع من انتشار الدرجات حول الدرجة المتوسطة، وعندها تصبح الأداة مناسبة للعينة من حيث درجة الصعوبة والسهولة؛
- أن تتمتع بحساسية عالية حيث أنه قد يتوفر في اختبار ما الثبات والصدق والموضوعية والشمول ولكنها غير حساسة، بمعنى أنها غير مناسبة لما نقيسه تحت الظروف الراهنة للقياس، فاختبارات الذكاء المخصصة للموهوبين والعباقرة من الأطفال رغم تمتعها بكل الخصائص الجيدة إلا أنها لا تصلح لاستخدامها في ذكاء الأطفال العاديين.

1. الموضوعية:

من أهم صفات الاختبار الجيد أن يكون موضوعياً في قياسه للنواحي التي أعد لقياسها، ويمكن تحقيق الموضوعية عن طريق: فهم الطالب لأهداف الاختبار والتعليمات فهماً جيداً كما يريدها واضع الاختبار، وأن يكون هناك تفسير واحد للأسئلة والإجابات المطلوبة، بالإضافة إلى توفر الظروف المادية كالتهوية والإضاءة، وتوفر الظروف النفسية وتجنب القلق، ويعتبر الاختبار موضوعياً إذا أعطى الدرجة نفسها بغض النظر عن من يصححه.

2. الصدق:

يعتبر الصدق ضروري لأداة القياس لذا فهو شرط أساسي من شروط الاختبار التحصيلي الجيد.

والصدق كخاصية متقدمة على الثبات من بين الطرق المعتمدة في التأكد من صلاحية الاختبار. فالصدق هو أن يقيس الاختبار ما وضع لقياسه فعلا، فالاختبار يكون صادقا إذا قاس الوظيفة التي وضع لأجلها ولا يقيس شيئا آخر بدلا منها أو وظيفة إضافية.

أنواع الصدق:

المطلع على البحوث التربوية والنفسية والاجتماعية يجد العديد من التسميات المتعلقة بصدق الأدوات مثل الصدق الظاهري(السطحي)، الصدق المنطقي، الصدق الإحصائي، صدق المضمون (المحتوى)، الصدق التجريبي، الصدق العاملي، الصدق الذاتي، الصدق التنبؤي، الصدق التلازمي، الصدق التمايزي والتطابقي.

ولكن جرت محاولات لتصنيف أنواع الصدق فيما يلي:

أ- صدق وصفي: ويعتمد على الدراسة التمهيدية الأولية للاختبار لمعرفة صلاحيته للتجريب أم لا ويشمل على الأنواع التالية:

الصدق الفرضي (البناء أو التشييد أو المفهوم)- الصدق السطحي (الظاهري أو صدق المحكمين)

الصدق المنطقي (صدق المضمون أو المحتوى)

ب-الصدق الإحصائي: ويعتمد على تحليل نتائج الاختبار بعد تجربته ويشتمل على الأنواع التالية:

الصدق الذاتي- الصدق التجريبي (التلازمي والتنبؤي)- الصدق العاملي.

صنفت الجمعية الأمريكية لعلم النفس APA أنواع الصدق إلى 3 أنواع هي: صدق المحتوى- الصدق المرتبط بمحك – صدق التكوين الفرضي

صدق المحتوى: يدل صدق المحتوى على مدى تمثيل أسئلة الاختبار لمحتوى السمة أو الصفة المراد قياسها وتحديدها بحيث تمثل أسئلة الاختبار جميع مكونات الصفة المقاسة والاختبار يكون صادقا مادام يشمل جميع عناصر السمة أو القدرة المطلوب قياسها.

صدق المحتوى يتناسب بدرجة كبيرة مع الاختبارات التحصيلية حيث تكون الأسئلة ممثلة لجميع عناصر المحتوى أفضل تمثيل، أما في حالة الاختبارات السيكولوجية فإن الأمر يختلف بعض الشيء حيث يتطلب الأمر تحليل أعمق لمجال القياس (عقلي أو انفعالي أو نفس حركي) ويمكن ذلك عن طريق أسلوب تحليل المحتوى.

طرق تقدير صدق المحتوى:

- في حالة الاختبارات السيكولوجية: تعتمد معظم أساليب تقدير صدق المحتوى على أحكام وآراء الخبراء أي تتعلق بمدى التناظر بين مفردات الاختبار ومحتوى الصفة المراد قياسها. يتم عرض الاختبار على مجموعة من المحكمين المتخصصين لا يقل عددهم عن 5 بغرض الكشف عن مدى الاتفاق في تقديراتهم ويفضل ألا تقل درجة الاتفاق بينهم على كل بند من بنود الاختبار 80%.
- في حالة الاختبارات التحصيلية: تعتمد على جدول مواصفات الاختبارات ويمكن حساب صدق المضمون من حساب معامل الارتباط بين درجات الأفراد في كل سؤال ودرجاتهم في الاختبار ككل.

الصدق المرتبط بمحك: (الصدق التجريبي–الصدق العلمي– صدق الوقائع الخارجية)

يقصد بالمحك أساس خارجي مستقل للحكم على مدى صلاحية الاختبار لما يقيسه أي هو اختبار للاختبار. صورة الصدق المرتبط بمحك صورتان هما: الصدق التنبؤي والصدق التلازمي، ويتم حساب الصدق التلازمي من خلال حساب معامل ارتباط بيرسون بين درجات الأفراد على الاختبار ودرجاتهم على محك خارجي ثبت صدقه في نفس الوقت أي لا يوجد فاصل زمني كبير بين أداء الأفراد على الاختبار ودرجاتهم على المحك ، أما الصدق التنبؤي فيدل على مدى كفاءة درجات اختبار ما في التنبؤ بسلوك الأفراد في موقف لاحق والذي يقاس باختبار محك، ويتم حساب الصدق التنبؤي من خلال حساب معامل الارتباط بين درجات مجموعة من الأفراد على اختبار ما ودرجاتهم على اختبار آخر (محك) يتم تطبيقه بعد فاصل زمني لا يقل عن ستة أشهر وقد يمتد لعدة سنوات من تطبيق الاختبار.

الصدق التنبؤي يستخدم في أغراض التنبؤ طويل المدى، لذا فالمعلومات التي يوفرها صدق الاختبار تستخدم في انتقاء الأفراد أو تصنيفهم أو توجيههم تربويا لنوع التعلم (الشعبة) وكذلك التوجيه المهني.

صدق التكوين الفرضي: (صدق المفهوم– صدق البناء– الصدق المنطقي)

ويقصد به مدى قياس الاختبار لتكوين فرض قائم على تصور نظري أو على نظرية سيكولوجية وهو في هذه الحالة يعد محاولة لإثبات صحة النظرية السيكولوجية أو التصور النظري الذي وضع على أساسه هذا الاختبار. هذا النوع من الصدق له أهميته في بناء الاختبارات النفسية عامة والعقلية بصفة خاصة.

طرق حساب معامل صدق الاختبار:

- طريقة استطلاع آراء المحكمين.
- طريقة معاملات الارتباط.
- طريقة المقارنة الطرفية: وتقوم في جوهرها على مقارنة متوسط درجات الأقوياء (الثلث الأعلى للدرجات) بمتوسط درجات الضعفاء في الاختبار(الثلث الأدنى للدرجات) ثم حساب الدلالة الإحصائية للفرق بين المتوسطين، فإذا كانت هناك دلالة إحصائية واضحة بين متوسطي الأقوياء والضعفاء دل ذلك على صدق الاختبار.

كما نستطيع مقارنة متوسطي مجموعة الأقوياء في الاختبار والمحك ومقارنة متوسطي درجات الضعفاء في الاختبار والمحك فإذا لم تكن هناك دلالة إحصائية واضحة دل ذلك على صدق الاختبار.

طريقة التحليل العاملي: تتلخص طريقة حساب الصدق العاملي في النقاط التالية: إعداد بطارية من الاختبارات المراد تعيين معامل الصدق لها، تطبق الاختبارات على الأفراد ثم تحسب معاملات الارتباط البينية لمجموعة الاختبارات وتوضع في هيئة مصفوفة ارتباطية، تحلل مصفوفة معاملات الارتباط باستخدام إحدى البرامج الإحصائية مثل spss من اجل الوصول إلى تشبعات الاختبارات بالعوامل.

طريقة جداول التوقع: تعتمد هذه الطريقة على مقارنة التوزيع التكراري لدرجات الأفراد في الاختبار المطلوب وتعيين معامل صدقه بالتوزيع التكراري لدرجات الأفراد في المحك، ويتم تنظيم التكرارات والنسب المئوية المناظرة لها في جدول يسمى جدول التوقع ويساعد على تقدير مدى صدق الاختبار بالنسبة لكل مستوى من مستويات المحك الخارجي.

3. الثبات:

ويقصد بثبات الاختبار أن يعطي الاختبار نتائج ثابتة بغض النظر عن الزمان والمكان الذي يطبق فيه، أي أن النتائج لا تتأثر بزمان التطبيق أو المكان الذي يتم فيه.

ويشير مصطلح الثبات إلى الثقة الكبيرة في ثبات الاختبار لعلامات الطلاب حيث أن العلامة التي يحصل عليها الطالب (س) ستكون العلامة نفسها أو قريب من العلامة التي سيحصل عليها الطالب في المرات القادمة تحت نفس الظروف نفسها.

متى ما كانت درجات 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 مقياسا ثابتا. ولهذا فإن الثبات هو الاتساق والدقة في القياس.

وللثبات أشكال، منها: الثبات بطريقة التطبيق وإعادة التطبيق، والثبات بطريقة الصور المتكافئة، والثبات بطريقة الاتساق الداخلي.

- **طرق حساب الثبات:** يمكن تقدير معامل الثبات للاختبارات التحصيلية من خلال الطرق الآتية:

الثبات بطريقة التطبيق وإعادة التطبيق:

ويتم التحقق من الثبات بالإعادة من خلال تطبيق الاختبار على عينة من الطلبة، وبعد فترة يتم إعادة تطبيق الاختبار نفسه على العينة نفسها، وحساب معامل الارتباط بين علامات الطلبة على كل من الاختبارين القبلي والبعدي، فيكون معامل الارتباط هو قيمة ثبات الاختبار، وهذا يعني أنه كلما كان معامل الارتباط دالاً إحصائياً، يكون ثبات الاختبار عالياً.

الثبات بطريقة الصور المتكافئة:

ويتم التحقق من الثبات بالصورة المتكافئة من خلال إعداد صورتين متكافئتين من جميع الجوانب للاختبار، بحيث تتوافر فيهما المواصفات نفسها وعدد الأسئلة والصياغة والمحتوى ومستوى الصعوبة والأهداف والتعليمات والزمن، ثم تطبق الصورة الأولى من الاختبار، وبعد فترة زمنية، يتم تطبيق الصورة الثانية من الاختبار، وحساب معامل الارتباط بين علامات الطلبة على الصورتين المتكافئتين من الاختبار، ويكون معامل الارتباط هو قيمة معامل الثبات للاختبار.

الثبات بطريقة الاتساق الداخلي:

تختلف هذه الطريقة عن الطريقتين السابقتين في عدد مرات تطبيق الاختبار، حيث تتطلب الطريقتان السابقتان تطبيق الاختبار مرتين، فيما تتطلب هذه الطريقة تطبيق الاختبار مرة واحدة، ومن أشكال الثبات بطريقة الاتساق الداخلي: طريقة التجزئة النصفية، وطريقة كودر – ريتشاردسون، وطريقة ألفا كرونباخ.

4. **الشمولية:**

يتمتع الاختبار بصفة الشمولية إذا كانت بنوده الاختبارية تغطي جميع نقاط الموضوعات المقررة والتي يفترض أن يكون معد الاختبار قد قام بتنظيم تعلم طلابه لها، هذا بالإضافة إلى مراعاتها لمستويات الطلاب المختلفة، وعدم الشمول هو عدم تناول جميع الموضوعات وعدم مراعاة المستويات المختلفة للطلاب وذلك بسبب الحصول على نتائج لا تدل على واقع تحصيل الطلاب.

5. معامل الصعوبة والسهولة:

تعتبر دراسة الفقرات من حيث درجة الصعوبة عملية ضرورية وهامة في بناء الاختبار لأنها تساهم في الحكم على صلاحية الفقرة ومناسبتها لأغراض القياس، حيث يتم بناء على ذلك وضع الفقرة في مكانها المناسب من حيث الترتيب العام لفقرات الاختبار أو بنوده، وترتب عادة الفقرات في الاختبارات بطريقة واحدة حيث تكون الفقرة السهلة في بداية تصميم الاختبار وتكون الفقرة المتوسطة الصعوبة في الوسط وتكون الفقرة الصعبة في النهاية.

6. معامل التمييز: يشير معامل تمييز السؤال إلى مدى قدرة هذا السؤال على إبراز الفروق الفردية بين مستوى الكفايات، أي أنه يشير إلى درجة تمييز المفردة بين مرتفعي التحصيل الدراسي ومنخفضي التحصيل الدراسي من الطلاب بعد تطبيق الاختبار عليهم، فالطلاب مرتفعي التحصيل يجيبون عليه إجابة صحيحة. يرتبط معامل التمييز إلى درجة كبيرة بمعامل الصعوبة، فإذا كان الغرض من الاختبار هو أن يفرق بين القادرين من الطلاب وأولئك الأقل قدرة فإن السؤال المميز هو ما يقود إلى هذا الغرض، إذ أن مهمة معامل التمييز ينبغي أن تتمثل في تحديد مدى فاعلية سؤال ما في التمييز بين الطالب ذي القدرة العالية والطالب الضعيف بالقدر نفسه الذي يفرق الاختبار بينهما في الدرجة النهائية بصورة عامة.

7. فعالية البدائل (المشتتات):

المشتتات عبارة عن الخيارات الأخرى غير الصحيحة للسؤال ذي الاختيار من متعدد، ويفترض في المشتتات التالي:

- أن تكون جذابة وبالذات للأفراد الذين لا يعرفون الاختيار الصحيح، أي يتم اختيار أي مشتت من قبل مفحوص أو أكثر بنسبة لا تقل عن 5% من المفحوصين.
- المشتتات عبارة عن إجابة خاطئة، لذلك فالمشتت الجيد هو الذي يكون عدد المفحوصين الذين يختارونه من الفئة العليا أقل من عددهم في الفئة الدنيا أي أنه (يميز باتجاه معاكس لتمييز الفقرة). وكلما كانت قيمة المموه بالسالب كان هذا دليل على أن المشتت جيد.

8. القابلية للاستعمال: ويعني ذلك أن يتصف الاختبار بكل من الخصائص التالية:

- سهولة التصميم من حيث الوقت والجهد؛

وإمكانية التطبيق من حيث:

- وضوح التعليمات بالنسبة للطلاب والقائمين على التطبيق معا؛
- الوقت اللازم للتطبيق؛
- توافر الكفاية المطلوبة لدى القائمين على التطبيق؛
- سهولة تقدير الدرجات؛

– سهولة تفسير نتائج الاختبار؛

– تكلفته المالية في حدود الميزانية المتاحة.

9. الواقعية: في بناء الاختبار وتطبيقه وتصحيحه، يجب أن تراعي الظروف والإمكانات المتاحة، فعدد الأسئلة مثلا يجب أن يتناسب مع الزمن الاختبار وعمر التلميذ.