الفصل السابع

استخدام الاختبارات بصفة عامة
(التقديم - التقدير - التقويم)

* كتابات ومراجعة الفقرات
* ترتيب الفقرات
* اعداد تعليمات الاختبار
* تصحيح الاختبار
* التجهيز والتقديم
* تحليل الاختبار
* أساليب وطرق تفسير نتائج الاختبارات
* الصدق والثبات

obeikandl.com

obeikandl.com

عند تصميم فقرات الاختبار التحصيلي، يكون من الطبيعي والمتطلب تقديمها وتهيئتها للاستخدام. والأسلوب الأفضل هو إعداد مجموعة من الفقرات يوميا بينما يكون التعليم قائما في نفس الوقت وتكون المادة التعليمية ما زالت نشطة في الذهن. وعلى أية حال، يكون الإعداد المبكر ذا فوائد لوضع مجموعة من الفقرات في أثناء فترة مناقشة المادة التعليمية داخل حجرات الدراسة. كما أنه من المتطلب والمرغوب إعداد عدد من الفقرات يزيد عن العدد الذي تستوعبه خلايا جدول المواصفات، فقد ترى أثناء عمليات المراجعة حذف بعض منها أو سحبها جانبا لاستخدام آخر في وقت لاحق.

يجب أن تسجل كل فقرة اختبارية على ورقة أو على كارت مستقل، فهذا يسهل كثيرا واجب مراجعة الفقرات وترتيبها في الاختبار وإعدادها رهن الاستخدام. ويكون هذا الكارت بمثابة المرجع المناسب لتسجيل البيانات التحليلية الخاصة بالفقرة المسجلة عليه بعد اجراء تقويم فاعليته.

كتابة ومراجعة الفقرات

ان تصميم مجموعة الفقرات على فترات متناوبة بواسطة واضعها أو شخص آخر يمنح الفرصة لاجراءات وتعديلات في جانب جودة الاختبار وفاعليته، ويمكن على ضوء الأسئلة التالية اجراء تقويم دقيق للفقرات:

١- هل تقيس كل فقرة اختبارية معطى تعليميا هاما ورد في جدول المواصفات؟

يجب أن ترتبط كل فقرة اختبارية بواحدة من خلايا جدول المواصفات، فاذا

كانت الخلية التي تتصل بها الفقرة الاختبارية لوحظت على الكارت في نفس الوقت عند تصميم الفقرة، يكون الواجب اذن قراءة الفقرة ومراجعتها للاستدلال عن ملاءمتها، وعليه يكون التحليل النهائي موضحا مدى ارتباط واتصال كل فقرة اختبارية مباشرة بنوع ولون السلوك المحدد والمعين بواسطة المعطيات في قائمة المواصفات.

٢ـ هل يتناسب كل نوع من الفقرات مع المعطى التعليمي المعين والخاضع للقياس؟

هناك بعض المعطيات التعليمية يمكن قياسها بأي نوع من أنواع الفقرات المعروفة، ففي بعض الأحيان تفضل فقرات الاختيار من متعدد. واذا كانت المعطيات التعليمية تستدعي اجابات منشأة (امداد) تكون فقرات التكملة أو أسئلة المقال ملائمة للاستخدام، وعندما يكون هناك بديلان فقط متوفران يكون من المفيد استخدام فقرات الصواب والخطأ. واذا كانت المعطيات تستدعي سلسلة متصلة من العناصر تكون فقرات المقابلة والمزاوجة أكثر ملاءمة وعندئذ تكون المراجعة واعادتها أفضل أسلوب لتحقيق ما ذكرناه عن الأنواع المناسبة للأغراض المناسبة.

٣ـ هل تقدم كل فقرة واجبا واضحا دقيقا؟

يجب أن تكون المشكلة التي تقدمها الفقرة الاختبارية ـ بصرف النظر عن نوع الفقرة ـ واضحة دقيقة ولا لبس فيها لجميع المتعلمين، بحيث يكون فهمهم للواجب المطلوب واضحا ومناسبا للأداء الذي سيظهرون. أما تلك الفقرات الضعيفة في تحديد وتوضيح الواجب المتطلب، فلا يعود العيب عليها ولكن على واضعها، فيعوزه عندئذ المعرفة والمهارة والتدريب.

٤ـ هل صيغت الفقرات بأسلوب سهل واضح؟

بالاضافة الى التأكيد على مستوى القراءة وملاءمته لأعمار المجموعة المستخدمة للاختبار، وباستثناء المصطلحات الفنية كجزء من المشكلة، يجب أن تكون القواعد اللغوية المستخدمة في الصياغة سهلة بسيطة واضحة. فالجمل

القصيرة السهلة أفضل من الأخرى الطويلة المعقدة، ومراعاة ذلك يمكن تجنب الغموض واللبس، والعمل على معاونة القارىء الضعيف على عرض مستواه التحصيلي على نحو ملائم.

٥ـ هل الفقرات متحررة من التلميحات الغريبة والدخيلة؟

على الرغم من حرصنا على عدم تقصير المتعلم في الفقرة اذا حصل المعطى الخاضع للقياس فاننا حريصون أيضا على اجابته للفقرة اجابة صحيحة عندما لا يحصل المعطى المتطلب. ولذلك، فإن مراجعة الفقرات هي فرصة أخرى لحذف التلميحات التي تقود الى الاجابة الصحيحة مع بقاء عدم تحصيل المعطى المتطلب.

٦ـ هل الفقرات ذات مستوى مناسب من الصعوبة والسهولة؟[٢٦]

يجب أن تقابل صعوبة الفقرات، فقرات الاختبار القائم على المحك صعوبة الواجبات التعليمية المعينة في المعطيات التعليمية، وعليه يكون السؤال هو: «هل صعوبة الفقرة الاختبارية هي نفس الصعوبة المعينة في الواجب التعليمي؟»، وعليه يجب أن نأخذ في الاعتبار ملاءمات الواجب التعليمي لأعمار المتعلمين وقائمة المعطيات السلوكية.

وعند تقويم صعوبة الفقرات الاختبارية يجب أن نركز على السؤال: «ما مدى فاعلية هذه الفقرة في التمييز بين التحصيل العالي والمنخفض؟»، ويمكن القول ان بعضاً من الفقرات البسيطة والسهلة مع بداية الاختبار تعمل على التنشيط وحفز الهمم، وبعضا من الفقرات الصعبة عند نهاية الاختبار لتتحدى قدرات المتعلمين، والبقية الباقية بين هذه وتلك تكون على مستوى ٥٠٪ من الصعوبة.

٧ـ هل كل فقرة مستقلة، وهل الفقرات كمجموعة متحررة من التداخل؟

من الواجب مراعاة ألا تكون معرفة اجابة على فقرة تعتمد على معرفة اجابة

---

(٢٦) محمد رضا البغدادي (١٩٧٦ م) «مقارنة مدى فاعلية عدة طرق في تدريس وحدة الوراثة لطلاب دور المعلمين والمعلمات»، رسالة دكتوراه، مرجع سابق.

على فقرة أخرى. أي من الواجب أن تكون كل فقرة وحدة مستقلة والا أصبحت الاجابة على فقرة ما ايحاء مباشراً أو غير مباشر غير مرغوب فيهما ليكونا بين فقرات الاختبار.

٨ـ هل الفقرات التي يتضمنها الاختبار
تمثل تغطية مناسبة لجدول المواصفات ؟

قد يكون حذف أو اضافة بعض الفقرات عند مراجعة الاختبار عاملا غير مباشر للانحراف عن جدول المواصفات، وعليه فمن الواجب عند اجراء تعديل أو حذف أو اضافة فقرات جديدة أن يكون الاختيار النهائي للفقرات التي سيحتويها الاختبار بحيث تغطي ما جاء في خلايا جدول المواصفات.

بالاضافة الى الأسئلة السابقة التي يمكن اعتبارها **قواعداً** لتصميم الفقرات يمكن أن تكون أيضا معيارا لتقويم الفقرة وذلك **بإعداد قائمة مراجعة** للمعيار القائم على قواعد وأسس كل نوع من أنواع الفقرات الاختبارية.

ترتيب الفقرات الاختبارية

عند انتهاء مرحلة اختيار الفقرات التي سيتضمنها الاختبار يجب أن يكون هناك مراعاة لترتيب ورود هذه الفقرات في الاختبار، وفيما يلي بعض الأمور التي تنظم عملية ترتيب الفقرات:

١ـ يجب مراعاة تجميع الفقرات التي تقيس نفس المعطى التعليمي سويا.

تميل بعض أمور ومتطلبات الاستخدامات التربوية والاحصائية لنتائج الاختبار الى ترتيب فقرات الاختبار تبعا للمعطيات التعليمية الخاضعة للقياس والاختبار. فكل الفقرات التي تقيس معطى تعليمي تجمع سويا بعنوان مناسب. ويجب أن تكون العنونة تحت أحد التصنيفات الرئيسة مثل: (المعرفة- الفهم- التطبيق- . . .) وذات صياغة للأهداف التعليمية العامة («يعرف المصطلحات»، «يعرف الاسس والمبادى»، . . .) أو ذات صياغة للمعطيات التعليمية («يحدد المصطلحات»، «يكتب جملة مستخدما كل مصطلح»، . . .)، وسواء تم تجميع

الفقرات بمعايير عامة أو بمعطيات فإن ذلك يعتمد على المدى الواسع لنوع الاختبار. فعادة يكون وضع الاختبارات القائمة على المعايير على أساس المعايير العامة ذا دلالة جوهرية.

وخلاصة القول، ان تضمين الاختبار بالعناوين المناسبة يساعد المعلم على تحديد أنواع المعطيات التعليمية ذات الصعوبة، والعمل على تخطيط نظم تعليمية أخرى تصلح مع الأفراد والمجموعات كوسائل علاجية.

٢ـ يجب مراعاة تجميع الفقرات ذات النوع الواحد سويا

من المرغوب فيه تجميع كل فقرات الاختيار من متعدد في مجموعة مستقلة، وفقرات الاجابة الموجزة في مجموعة أخرى، وفقرات المقال في مجموعة أخرى، وهكذا. هذا الترتيب والتجميع الموحد يساعد كثيرا في وضع مجموعة واحدة من التعليمات لكل نوع من الفقرات. كما أن هذا الأمر يشارك بفاعلية في تقديم واجراء الاختبار حيث يستطيع المتعلم تحديد الأسلوب وطريقة الاجابة خلال كل قسم، بالاضافة الى المساهمة في تيسير عملية تقدير الدرجات وتحليل النتائج.

وفي حالة تعارض ترتيب الفقرات تبعا لنوعيتها مع ترتيبها تبعا للمعطيات التعليمية يكون اتباع ترتيبها تبعا للمعطيات أفضل كثيرا. علما بأن النوعين من الترتيب يمكن التكيف بهما وذلك لأن اختبارات التحصيل محدودة بعدد من أنواع الفقرات وتميل الفقرات التي تقيس معطى تعليميا معينا أن تكون من نوع واحد.

٣ـ يجب ترتيب الفقرات على نحو متصاعد في الصعوبة

من المرغوب فيه أن يبدأ الاختبار بالفقرات السهلة ثم تتدرج على نحو متصاعد من الصعوبة، فهذا الترتيب له دور تنشيطي مؤثر على المتعلمين ويمنع ويجنب المتعلم الضعيف من السقوط في هوة الفقرات الصعبة مبكرا. فاذا تم ترتيب الفقرات تبعا للمعطى التعليمي فمن الممكن ترتيب هذه المعطيات تبعا لصعوبتها على نحو متزايد من الصعوبة، وسوف يقابل هذا النظام تقريبا الوضع المتطلب للصعوبة المتزايدة.

إعداد تعليمات الاختبار

من المفضل أن تكون تعليمات اختبار التحصيل بسيطة موجزة واضحة متضمنة معارف تتصل بكل من:

١ـ الغرض من الاختبار

٢ـ الزمن المسموح به للأداء

٣ـ كيفية تسجيل الاجابات

والمثال التالي لتعليمات اختبار من نوع الاختيار من متعدد تشمل العناصر الثلاثة التي أشرنا اليها:

مثال

تعليمات:

*يقيس هذا الاختبار ما تعلمته ودرسته في الجزء الأول من المقرر، ونتائج هذا الاختبار سوف تستخدم في التوضيح والكشف عن المواضع والنقاط الصعبة والتي تساعد في استكمال دراسة المقرر بنجاح وكما ينبغي.

* يتألف الاختبار من (٩٠) تسعين فقرة من نوع الاختيار من متعدد، والزمن المسموح به ساعة ونصف الساعة فقط.

* تخير الاجابة الصحيحة لكل فقرة وضع دائرة حول رمز هذه الاجابة.

* تأكد من الاجابة عن كل الفقرات، فعدد الفقرات الصحيحة التي تجيب عنها مرتبط بالدرجة التي تستأهلها.

وعندما يكون الاختبار شاملا لنوعين أو أكثر من الفقرات الاختبارية، فمن المرغوب فيه أن تكون هناك تعليمات عامة للاختبار ككل، ثم تعليمات خاصة لكل نوع. وفي هذه الحالة يجب أن يتضمن نص التعليمات العامة على معارف عن الغرض والزمن ومآخذ التخمين، وتتضمن التعليمات الخاصة بكل نوع كيفية

تسجيل الاجابات، مع العلم بأن هناك بعض الفقرات تحتاج الى تعليمات خصوصية لكل فقرة.

أما في حالة استخدام ورقة اجابة مستقلة، فيجب مراعاة بعض الجهد والاتقان لتسجيل الاجابات، وفي حالة ما اذا كان المتعلم على غير ألفة باستخدام ورقة الاجابة المستقلة فيتعين اضافة مثال توضيحي لكيفية الاجابة وذلك في التعليمات.

وهناك أشكال عديدة لأوراق الاجابة المستقلة وعليه فان التعليمات ترتبط بنوعية هذه الأوراق وأشكالها وطريقة استخدامها وكيفية تقدير الدرجات من حيث التقدير اليدوي أو التقدير الآلي.

ويعتبر الشكل التالي (شكل ٦) نموذجا لورقة اجابة مستقلة من أفضل النماذج وأيسرها للمتعلم والمعلم.

بسم الله الرحمن الرحيم

وزارة.....

ادارة.....

مدرسة.....

الدرجة المستحقة

ورقة اجابة اختبار.....

رقم	الاجابات					رقم	الاجابات			
	أ	ب	ج	د			أ	ب	ج	د
١	()	()	()	()	٤٦		()	()	()	()
٢	()	()	()	()	٤٧		()	()	()	()
٣	()	()	()	()	٤٨		()	()	()	()
٤	()	()	()	()	٤٩		()	()	()	()
٥	()	()	()	()	٥٠		()	()	()	()
٤٥					٩٠					

شكل (٦)

نموذج ورقة اجابة مستقلة

ويطلب من المتعلم في هذا النوع من أوراق الاجابة أن يظلل أسفل رمز الاجابة أمام السؤال المعين التي يرى أنها صحيحة على السؤال.

وتعتبر طريقة التصحيح هنا عملية يسرة حيث يمكن استخدام المفتاح المثقب الذي يعده المعلم مسبقا، وذلك بتثقيب ورقة اجابة في الدوائر الصحيحة أسفل رموز الاجابة مقابل كل سؤال.

تصحيح الاختبار [1]

من العوامل الهامة التي تؤثر في الاختبار التحصيلي طريقة التصحيح، أو بعبارة أخرى القانون المستخدم في التصحيح ووزن كل بند من بنوده، ونوع مفتاح التصحيح المستخدم. لهذا كانت العوامل المؤثرة في الاختبار ذات ضرورة وأهمية يجب مراعاتها عند تصميم الاختبار. كذلك يجب أن يفكر واضع الاختبار في كيفية تسجيل الاجابات ثم كيفية التصحيح، هل سيتم بالآلة أم يدوياً.

ومن أهم ما يجب أن يتخذه واضع الاعتبار الموضوعي في حسبانه ما يتعلق بتصحيح الاختبار، هل ستكون الدرجة هي عدد الاجابات الصحيحة بالنسبة للعدد الكلي؟ ، أم لا بد من استخدام قانون لتصحيح الدرجة الخام من أثر التخمين؟ [2].

وهناك قانون يمكن استخدامه للتصحيح من أثر التخمين ثبتت فاعليته في هذا الصدد، وهو:

$$د = ص - \frac{خ}{ب - ١}$$

حيث أن:

د = الدرجة المستحقة.

ص = عدد الاجابات الصحيحة.

---

(١) رمزية الغريب (١٩٧٠) التقويم والقياس النفسي والتربوي. القاهرة: مكتبة الأنجلو المصرية

(٢) محمد رضا البغدادي (١٩٧٦ م) مرجع سابق.

خ = عدد الاجابات الخاطئة.

ب = عدد البدائل.

والقانون العام الذي اشتق منه القانون السابق هو:

$$س = ص - \frac{ك \times خ}{ن - ك}$$

حيث أن:

س = الدرجة المستحقة.

ص = عدد الاجابات الصحيحة.

خ = عدد الاجابات الخاطئة.

ن = عدد الاجابات التي تختار منها الاجابة الصحيحة (عدد البدائل).

ك = عدد الاجابات المختارة من كل سؤال (علمًا بأنها إجابة واحدة).

يلاحظ أن هذا القانون يفترض أن كل إجابة خاطئة نتيجة للتخمين، وأن الاجابات الخاطئة تتساوى في احتمال اختيارها بالتخمين، ولهذا كان قانون الصدفة يمكن تطبيقه في هذا الموقف.

والجدير بالذكر أن هذه الافتراضات لا يمكن تحقيقها تماماً، لأن الاجابات الخاطئة محتمل أن تكون نتيجة للجهل بمعارف معينة وليس نتيجة الخطأ في التعليم، كما أن واضع مفردات الاختبار لن يوفق عملياً في جعل جميع الاجابات الخاطئة ذات صياغة تجذب المتعلم بنفس القدر.

ومن الواضح أنه كلما زاد عدد البدائل أو فرص الاختيار كلما قلّ أثر التخمين، ولهذا ينصح البعض بعدم استخدام معادلة التصحيح من أثر التخمين اذا زادت فرص الاختيار عن خمسة بدائل، ومع ذلك فان بعض مصممي الاختبارات يرون ضرورة من هذا التصحيح مهما كان عدد هذه البدائل.

ويرى كثير من العلماء أمثال (رخ Ruch ، دي جراف Degraff و ود Wood، وغيرهم) أن هناك ما يبرر استخدام طريقة التصحيح من أثر

التخمين لأنه ولو أن التصحيح لا يزيد من معامل ثبات الاختبار الا أنه يزيد قليلا من معامل الصدق [1].

ونرى أنه اذا أراد المعلم أن يقوم بعمل اختبار ما لاستخدامه في نطاق ضيق محدود فليست هناك حاجة الى استخدام قانون التصحيح من أثر التخمين لأن ما سيحصل عليه من فوائد في تصحيح الدرجات الخام قليل لا يتناسب مع ما يبذل من مجهود.

ويلاحظ أن هناك نوعاً من الاختبارات الموضوعية لا يصلح معها استخدام قانون التصحيح من أثر التخمين، وهو اختبار المطابقة أو التطابق. ويرى الكثير أن استخدام قانون التخمين يميل الى المبالغة في أثر التخمين، كذلك من الواضح أن التصحيح من أثر التخمين لا يمكن أن يستبعد هذا الأثر تماماً لأننا لا نستطيع أن نستخدم قانوناً يقابل تماماً أثر الحظ في إجابات بعض المتعلمين.

وتختلف طرق التصحيح من اختبار لآخر حسب القانون الذي يستخدم في تقدير درجة الاختبار، وهناك عدة طرق منها:

أ- اذا قدرت الدرجة على أساس القانون:

د = عدد الاجابات الصحيحة

كانت طريقة التصحيح قاصرة على وضع علامة على الاجابة الصحيحة، ثم تعديد هذه العلامات دون أن تؤخذ الاجابات الخاطئة والمتروكة في الاعتبار. واذا استخدمت ورقة اجابة مستقلة كان من السهل عدّ الاجابات الصحيحة مع استخدام مفتاح التصحيح.

ب- اذا قدرت الدرجة على أساس القانون:

---

(1) Traxler, A. (1930) «Administering and Scoring the Objective Tests», In **Educational Measurement,** ed., by Lindquisi. Vol. 32. PP, 142 - 157.

$$د = ص - \frac{خ}{ب - ١}$$

كان لا بد من عدّ الاجابات الصحيحة، والاجابات الخاطئة، ومن المستحسن أن تعد الاجابات المتروكة أيضاً.

وفي حالة عدم التصحيح في الاختبار نفسه يفضل عمل ثلاثةأعمدة للتصحيح (ص، خ، م «متروك»)، وذلك حتى تسهل عملية عدّ الاجابات بأنواعها الثلاث. ويلاحظ أن قسمة الاجابات الخاطئة على العدد المناسب للحصول على نسبة الاجابات الخاطئة (ب- ١) التي يمكن خصمها من الدرجة لتصحيح أثر التخمين عملية شاقة، ولذلك أعدت قوائم يمكن الاستعانة بها في هذا الصدد:

جـ اذا كان القانون المستخدم للتصحيح هو:

$$د = ص \times + \frac{م}{ب}$$

حيث أن: د = الدرجة المستحقة

ص = عدد الاجابات الصحيحة.

م = عدد الاجابات المتروكة.

ب = عدد البدائل.

كان من الممكن هنا أن يقوم المتعلمون بأنفسهم بالخطوات المبدئية في التصحيح، حيث يطلب منهم قبل جمع أوراق إجابة الاختبار أن يعد كل منهم الاجابات التي تركها ويسجل العدد في المكان المخصص لذلك، ويحسن أن تعد هذه الاجابات بواسطة متعلم يراجع عليه آخر، وهكذا.

هناك العديد من مفاتيح التصحيح، وسوف نتناول بعضاً منها، وهي الأنواع الشائعة وأهم هذه المفاتيح:

١ـ المفتاح ذو المروحة [١]

هذا المفتاح هو أكثر شيوعاً في تصحيح الاختبار من كراسة الاختبار الأصلية، وقد اشتق اسمه من امكانية تطبيق المفتاح على حسب أقسامه الرأسية فيصبح في شكل المروحة، ويلاحظ أن بالمفتاح قوائم بدرجات خام حولت الى درجات معيارية حتى تكون قابلة للمقارنة.

٢ـ المفتاح الشريطي

ويشبه مفتاح التصحيح السابق، ولا يختلف عنه الا في كون التصحيح الخاص بكل صفحة من صفحات الاختبار قد كتبت على حدة، كما أن المفتاح قد كتب على ورق مقوى وليس على الورق العادي.

٣ـ المفتاح المثقب

في مثل هذا المفتاح تعمل ثقوب لبيان الكلمات أو الرموز أو الجمل الممثلة للاجابة في مفردات التكملة وقد كان هذا النوع شائعاً في الاستخدام حين كانت اختبارات التكملة سائدة، ولكنها وقد قلّ استخدامها حالياً بعد أن زاد انتشار استخدام اختبارات الاختيار من متعدد، وعليه فيعد المفتاح المثقب بثقيب ورقة الاجابة المستقلة (ويفضل أن تكون مطبوعة على الورق المقوى) لبيان رموز الاجابات الصحيحة أمام أرقام الأسئلة.

٤ـ المفتاح الشفاف

يتميز هذا المفتاح بأنه مصنوع من رقائق البلاستيك الشفاف أو الورق الشفاف، والشفافية تجعل المفتاح معيناً على سرعة التصحيح وتمييز المفردات التي قد تصحح مرتين، ولهذا يكثر استخدامه لفاعليته في تصحيح الكثير من الاختبارات الموضوعية.

---

(١) رمزية الغريب (١٩٧٠)، مرجع سابق.

٥ـ مفتاح الكربون

تصاحب هذا المفتاح ورقة الاجابة، حيث تحدد أماكن الاجابات الصحيحة على ورقة مستقلة تلصق أطرافها في ظهر ورقة الاجابة بحيث تكون مستترة بالنسبة للمتعلم، ويطلى ظهر ورقة الاجابة بطلاء كربوني بحيث يترك أثراً لأية كتابة أو علامة تسجل على ورقة الاجابة.

وقد ظهر حالياً نوع من الورق (للكتابة) يستخدم كثيراً في الدوائر الحكومية يغني عن استخدام أوراق الكربون، ويعرف بالورق «المكربن» ظهر الصفحة مزودة بمادة كربونية عديمة اللون تظهر أية علامة على وجه الورقة التالية لها.

بعد اتمام الاجابة عن الاختيار يتناولها المصحح، وينزع عنها مفتاح الاجابة ويقوم بتعديد الاجابات الصحيحة ويسجلها. هذا النوع من المفاتيح سريع الا أنه كثير التكاليف ولا يستخدم الا حين توجد الامكانات المادية الضرورية.

٦ـ المفتاح المطبوع على ورقة الاجابة

يطبع المفتاح على ورقة الاجابة بعد أن ينتهي المتعلمون من الاجابة، وقد جرّب هذا المفتاح مع الآلاف من المتعلمين دفعة واحدة وثبت أنه يوفر ٤٢٪ من الوقت المستخدم في التصحيح.

٧ـ المفتاح الآلي

يمكن استخدام المفتاح المثقب في التصحيح بالآلات اذا أعدت أوراق الاجابة للتصحيح الآلي، واذا اتفقت ثقوب أوراق الاجابة مع الآلة. وتكون وظيفة المفتاح الآلي هو تقسيم نقط الاتصال الى دائرتين احداهما للاجابات الخاطئة.

التجهيز والتقديم

عندما يطبع الاختبار في صورته النهائية، يجب أن تكون فقراته موزعة على كل صفحة، وبين الأسطر مسافة مناسبة تسمح للمتعلم بالقراءة الواضحة السهلة وللمعلم بالتقدير المريح. فاذا كانت هناك فقرات الاختيار من متعدد مثلاً، يجب أن

تسجل البدائل في قائمة أسفل الدعامة (كما أوضحنا قبل ذلك)، ويراعى أن تكون أجزاء الفقرة الواحدة في صفحة واحدة، أما في حالة نوع الفقرة المركبة فيمكن أن تكون المادة التمهيدية في بداية الصفحة أو على صفحة مستقلة.

واذا كان التصحيح سيتم على نفس الاختبار، فمن الواجب مراعاة أن يكون تسجيل الاجابات في ناحية واحدة موحدة في جميع فقرات الاختبار، فهذا أيضاً ييسر التقدير، ويوحد النظام للمتعلم أما اذا كانت هناك أوراق إجابة مستقلة والاختبار نفسه يستخدم مع أكثر من مجموعة من المتعلمين فمن الضروري التنبيه على المتعلمين بعدم ترك أية علامة في كراسة الاختبار، ويراعى أيضاً طباعة عدد أكبر ـ للاحتياط ـ من كراسات الاختبار.

يعد الكثير من الاختبارات التحصيلية في صورة التطبيق مطبوعاً على آلات النسخ أو آلات السحب العادية، أو على آلات التصوير أو بعمليات الأوفست. ولما كانت آلات السحب هي الأنسب والأقل تكلفة للعديد من الأغراض، فان الرسوم والمخططات والأشكال تتطلب أسلوباً آخر في الطباعة بدلاً من آلات السحب حرصاً على وضوح الرسوم والاشكال في جميع النسخ.

ومن الضروري عند استخدام اسلوب آلات السحب أو آلات التصوير مراجعة الأوراق الأصول قبل السحب أو التصوير للتحرر من الأخطاء التي قد لا تكتشف الا بعد توزيع الاختبار أو الانتهاء من تجميعه.

ان تقديم الاختبار المعد إعداداً دقيقاً يتطلب ظروفاً للأداء مناسبة، فإبعاد الضوضاء ومسببات المقاطعة، وترتيب وتنظيم جلسة المتعلمين في مسافات مناسبة منعاً لحالات الغش ضرورات يجب مراعاتها.

كما أن التعليمات المكتوبة يجب أن تكون واضحة بما فيه الكفاية بحيث تجعل الاختبار يقدم نفسه، ولكن يتطلب في بعض الأحوال اعطاء التعليمات الشفوية، وقد يكون من المناسب والمفيد استخدام السبورة الطباشيرية مع المتعلمين صغار السن. وبصفة خاصة يجب التأكيد تماماً من معرفة جميع المتعلمين لما سوف يقومون به

ثم توفير جميع السبل والوسائل الممكنة لأداء مريح بعيد عن الازعاج والارباك.

يكون تقدير الدرجات عملاً سهلاً ميسوراً اذا سجلت الاجابات في جانب واحد موحد من صفحات الاختبار، وبذلك يكون عملاً روتينياً لعد الاجابات الصحيحة ثم تسجيلها في عمود الاجابة لكل متعلم. وفي حالة استخدام ورقة اجابة مستقلة يكون من الأفضل تثقيب الرموز الدالة على الاجابات الصحيحة على نسخة واحدة من أوراق الاجابة واستخدام هذا المفتاح المثقب في عد الاجابات الصحيحة التي تظهر علاماتها أسفل الثقوب. وفي حالة عدم ظهور علامة المتعلم أسفل الثقوب يكون من واجب المعلم القائم بالتصحيح وضع علامة حمراء على هذا الثقب حتى يوضح للمتعلم موضع الخطأ، أما في حالة استخدام التصحيح الآلي فيكون من السهل والميسور توفير الجهد على المعلم.

تحليل الاختبار

أولاً: تحليل الفقرات الاختبارية المعتمدة على المعايير

يكون من المتطلب بعد تقدير درجات الاختبار تقويم فاعلية الفقرات، ويتم هذا التقويم بدراسة اجابات المتعلمين على كل فقرة تضمنها الاختبار، ويعرف هذا الاجراء بتحليل الفقرات، وهو يزودنا بالمعارف المتصلة بكيفية توظيف كل فقرة من فقرات الاختبار. وسوف نتناول تحليل الفقرات التي تقوم على المعايير والتي تصمم للتمييز بين المتعلمين.

يوفر تحليل الفقرات الاختبارية المعتمدة على المعايير على المعارف الآتية:

١ـ صعوبة الفقرة.
٢ـ قوة الفقرة في التمييز.
٣ـ فعالية كل بديل.

وتدلنا المعارف ما اذا كانت الفقرة سهلة جداً أم صعبة جداً، وكيف تميز بين التقديرات العالية والمنخفضة في الاختبار، وما اذا كان قد تم توظيف كل بديل من

البدائل كما هو متطلب له. كما أن بيانات تحليل الفقرة تساعد بدورها في ضبط وكشف الأخطاء الفنية وتعطي معارف أكثر لتحقيق فاعلية فقرات الاختبار.

وهناك عدد من الخطوات تعتبر ذات فاعلية، وسوف يكون مثالنا هنا على اختبار طبق على عينة من المتعلمين والبيانات المعروضة مأخوذة عن (٣٢) اثنتين وثلاثين ورقة اجابة [١]:

١ـ رتب أوراق الاجابة (٣٢ ورقة اجابة) ترتيباً تنازلياً تبعاً للدرجات.

٢ـ تخيرَّ ثلث عدد الأوراق تقريباً من التي حصلت على درجات عالية، وتسمى هذه المجموعة «المجموعة العليا»، ثم تخيرَّ نفس العدد من الأوراق الأخرى التي حصلت على درجات منخفضة وتسمى هذه المجموعة «المجموعة المنخفضة». وعليه تكون المجموعة الوسط تتألف من (١٢) ورقة اجابة. وسوف نستخدم درجات المجموعتين العليا والدنيا فقط.

٣ـ بالنسبة لكل فقرة، احسب عدد المتعلمين في المجموعة العليا الذين تخيَّروا كل بديل (علمًا بأن البدائل الواردة تبعاً لكل فقرة خمسة بدائل)، وكرر نفس العمل مع المتعلمين في المجموعة المنخفضة.

٤ـ سجِّل هذه البيانات في ورقة منفصلة بالنسبة لجميع فقرات الاختبار، كما يلي:

الفقرة رقم (١)	البدائل				
	أ	ب*	جـ	د	هـ
المجموعة العليا (ن ١ = ١٠)	صفر	٦	٣	١	صفر
المجموعة المنخفضة (ن ٢ = ١٠)	٣	٢	٢	٣	صفر

٥ـ احسب صعوبة الفقرة، وذلك بتقدير النسبة المئوية للمتعلمين الذين

---

(١) Throndike, R.L., ed., (1971) **Educational Measurement.** Washington, D.C.,: American Council on Education.

أجابوا على الفقرة اجابة صحيحة، وذلك بجمع عدد المتعلمين الذين أجابوا على الفقرة اجابة صحيحة في كل من المجموعتين العليا والمنخفضة، فيكون كما في مثالنا (٦ + ٢ = ٨) ، علمًا بأن ن = ١٠ + ١٠ = ٢٠، ثم اقسم ٨ ÷ ٢٠ × ١٠٠ يكون التالي:

معامل صعوبة الفقرة رقم (١) = ٢٠ / ٨ × ١٠٠ = ٤٠٪

وعليه تكون المعادلة المستخدمة هي:

$$\text{س ع} = \frac{\text{ص}}{\text{ن}} \times ١٠٠$$

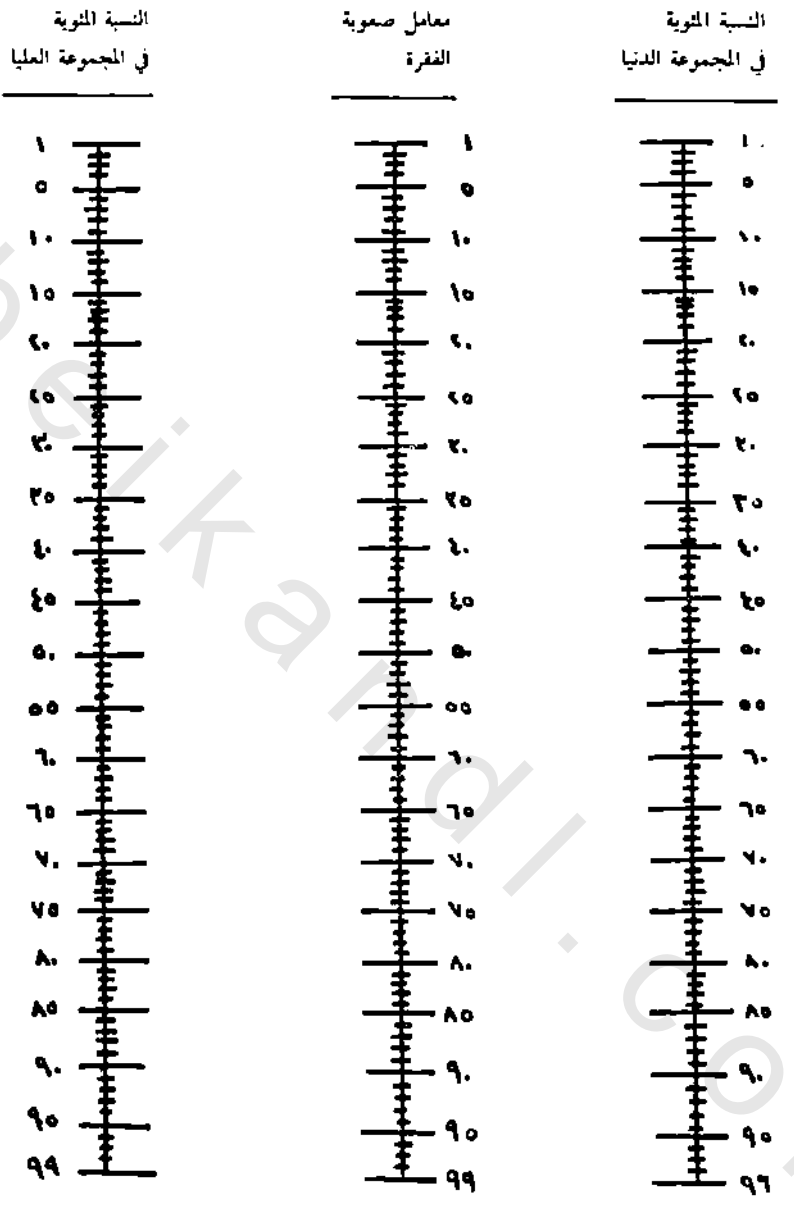
حيث أن س ع = معامل صعوبة الفقرة = (١ـ معامل سهولة الفقرة)

ص = عدد الاجابات الصحيحة على الفقرة (أو النسبة المئوية)

ن = العدد الكلي (مجموع أوراق المجموعة العليا + المجموعة المنخفضة)

ويمكن الاستعانة بالقائمة الموضحة في شكل (٧) بعد تحديد عدد أو النسبة المئوية للمتعلمين الذين أجابوا اجابة صحيحة على الفقرة في المجموعة العليا، وعدد أو النسبة المئوية للمتعلمين الذين أجابوا على نفس الفقرة في المجموعة المنخفضة، على أن يكون عدد المجموعتين متماثلاً.

وطريقة استخدامها للحصول على معامل صعوبة الفقرة الاختبارية سهلاً مع وجود البيانات السابقة، وذلك بتحديد نقطة تعين النسبة المئوية للمتعلمين الذين أجابوا اجابة صحيحة على الفقرة في المجموعة المنخفضة، ونقطة تعين النسبة المئوية للمتعلمين الذين أجابوا اجابة صحيحة على نفس الفقرة في المجموعة العليا، ثم صل خطاً مستقيمًا بين النقطتين ليقطع خط العمود الأوسط، فتكون نقطة التقاطع دالة على معامل صعوبة الفقرة الاختبارية المعينة.



شكل (٧)

طريقة حساب معامل صعوبة فقرة اختبارية،
بمعلومية النسب المئوية للمتعلمين الذين أجابوا على الفقرة إجابة صحيحة

٦ـ احسب قوة تمييز كل فقرة، وذلك بمقارنة عدد المتعلمين في المجموعة العليا والمجموعة المنخفضة الذين أجابوا اجابة صحيحة على الفقرة ، ولعلك تذكر في مثالنا السابق أن هذا الفرق هو (٦ـ ٢ = ٤)، وهذا يبين أن التمييز ايجابي.

وعليه فان حساب معامل التمييز يتم بطرح عدد الذين أجابوا اجابة صحيحة على الفقرة في المجموعة الضعيفة من الآخرين في المجموعة العليا ثم قسمة الناتج على عدد أفراد المجموعة الواحدة (علمًا بأن العدد متساو) ، فتكون المعادلة:

$$\text{م ت} = \frac{\text{ص ١} - \text{ص ٢}}{\text{ن}}$$ حيث أن:

م ت = معامل التمييز.

ص ١ = عدد الاجابات الصحيحة في المجموعة العليا.

ص ٢ = عدد الاجابات الصحيحة في المجموعة المنخفضة.

أي أن معامل التمييز = $\frac{٦ - ٢}{١٠} = ٠,٤$.

وعندما يصل معامل التمييز الى الواحد الصحيح فان هذا يعني مستوى عالياً من التمييز، حيث أن كل المتعلمين في المجموعة العليا أجابت اجابة صحيحة، بينما لم يجب أحد في المجموعة الدنيا (المنخفضة) وعندما يصل معامل التمييز الى الصفر فان هذا يعني عدم تمييز للفقرة، لأن كل المتعلمين في كلتا المجموعتين أجابوا إجابات صحيحة.

وعندما يصل معامل التمييز الى رقم باشارة سالبة فان هذا يعني ان عدد المتعلمين في المجموعة الضعيفة كان أكثر من الآخرين في المجموعة العليا.

٧ـ قدِّر فاعلية البدائل، وذلك بمقارنة عدد المتعلمين في المجموعة العليا بعدد المتعلمين في المجموعة الضعيفة الذين تخيروا كل بديل خاطىء فالمشتت الجيد سوف يجذب عدداً أكبر من المجموعة الضعيفة عن المجموعة العليا.

وعليه ففي الخطوة الرابعة السابقة يلاحظ البدائل (أ)، (د) ذات توظيف فعّال، بينما البديل (جـ) قد جذب عدداً أكبر من متعلمي المجموعة العليا، والبديل (هـ) غير ذي فاعلية. وعليه فان تحليلاً مثل التحليل السابق له فائدته في تقويم كل فقرة اختبارية، ويساعد كثيراً في تحسين الفقرات ورفع كفاءتها وفاعليتها.

ثانياً: تحليل فقرات الاختبار القائمة على المحك

لما كان الغرض الذي تصمم من أجله فقرات الاختبارات القائمة على المحك هو التعرف على وتوصيف الواجبات التي يستطيع المتعلم تأديتها والتي لا يستطيع تأديتها، فان هذا أسلوباً قد يكون أفضل من التمييز بين المتعلمين وأفضل من المعاملات التقليدية من صعوبة الفقرات، وقوة التمييز التي تعتبر حالياً ذات قيم قليلة. فمثلاً مجموعة الفقرات لاختبار محك للحفظ يمكن الاجابة عنها من جميع المتعلمين اجابة صحيحة (قوة تمييز صفرية) وتكون في نفس الوقت فقرات ذات فاعلية. أما اذا كانت الفقرات تقابل تماماً معطى تعليمياً هاماً، فان النتائج ستدل على أن هناك معطى تعليمياً هاماً قد تحقق من جميع المتعلمين ، وهذا يعتبر ذا قيمة معرفية لتوصيف أنواع الواجبات التي يمكن أن تؤدي وحذف الأخرى التي تشوّه التعلم.

يمكن تقدير صعوبة فقرة الاختبار القائم على المحك بواسطة الواجب التعليمي الذي يخطط ويخضع بعدئذ للقياس، فاذا كان الواجب سهلاً فان الفقرة ستكون سهلة أيضاً، واذا كان الواجب صعباً فان الفقرة ستكون حينئذ صعبة، ولن تكون هناك محاولات لحذف الفقرات السهلة أو لتعديل الأخرى الصعبة حتى يمكن الحصول على مدى منتشر من الدرجات الخام.

وحتى يمكن قياس فاعلية الفقرة ليقوم المعلم بتطبيق الاختبار قبل الفترة التعليمية (اختبار قبلي) وبعدها (اختبار بعدي)، والفقرة الفعالة سوف يجيب عنها عدد كبير من المتعلمين بعد الفترة التعليمية أكثر منها قبل الفترة التعليمية. وعليه

يمكن حساب معامل الحساسية للأثر التعلمي من المعادلة [1] :

$$\text{س} = \frac{\text{ن ٢} - \text{ن ١}}{\text{ن}}$$

حيث أن: س = معامل الحساسية للأثر التعلمي

ن١ = عدد المتعلمين الذين أجابوا على الفقرة بعد الفترة التعليمية اجابة صحيحة.

ن٢ = عدد المتعلمين الذي أجابوا على الفقرة قبل الفترة التعليمية اجابة صحيحة.

ن = عدد المتعلمين جميعاً الذين أجابوا على الفقرة في كلتا الحالتين.

وبتطبيق المعادلة على فقرة ما أجاب عليها كل المتعلمين (ن = ٣٢) اجابة خاطئة قبل الفترة التعليمية، وأجاب عليها كل المتعلمين (ن = ٣٢) اجابة صحيحة بعد الفترة التعليمية.

تكون النتيجة هي : $\text{س} = \frac{\text{٣٢} - \text{صفر}}{\text{صفر}} = ١,٠٠$

وهذا هو الحد الأقصى لمعامل الحساسية والذي يجب أن يتراوح بين الصفر والواحد الصحيح.

وهناك بعض الحدود يجب مراعاتها عند استخدام معامل الحساسية، منها:

١ـ يجب أن يقدم المعلم الاختبار مرتين لحساب المعامل.

٢ـ يمكن ارجاع المعامل الى كل من الفاعلية (فاعلية الفقرة) أو عدم فاعلية التعليم.

٣ـ اجابات المتعلمين على الفقرات بعد الفترة التعليمية يمكن أن تتأثر وخاصة

---

(١) Kryspin, W. and J.T. Feldhusen (1974) **Developing Classroom Tests.** Minneapolis: Burgess publishing Co.

أن الأختبار المطبق هو نفسه، وفي حالة تطبيقه بعد فترة قصيرة تسمح بتذكر بعض فقراته. وبالرغم من هذه الحدود فان معامل الحساسية له قيمته المفيدة في تقويم فاعلية الفقرات وبخاصة في اختبارات الحفظ.

أساليب وطرق تفسير نتائج الاختبارات

بعد الانتهاء من تقدير درجات الاختبار التحصيلي، يكون من الواجب عندئذ تنظيم هذه النتائج بأسلوب يسمح لها أن تصير جاهزة ومعدة للتفسير. ان الكيفية التي تنظم بها النتائج تقوم على ضوئها أساليب التفسير، فاذا كنا سنستخدم نتائج الاختبار لوصف أنواع الواجبات التي استطاع المتعلم تأديتها (تفسير المقاييس المعتمدة على المحك) يكون العرض والتحليل تفصيليا، أما اذا كانت الرغبة مجرد الاشارة لوضع أو علاقة المتعلم بالمجموعة (تفسير المقاييس المعتمدة على المعايير) يكون المطلوب هو مجرد تعديد الدرجات النهائية للاختبار.

وعليه سوف يدور نقاشنا في هذا الجزء عن كل نوع من التفسير والتي نفترض ان الاختبار قد صمم أساسا لطريقة التفسير التي ستستخدم.

تفسير الاختبارات المعتمدة على المحك

يتركز الانتباه والعمل في كلا التصميم والتفسير للاختبارات المعتمدة على المحك على الأهداف السلوكية المتطلبة للقياس.

ولما كانت الاختبارات المعتمدة على المحك تستخدم كثيرا في قياس الحفظ، فإن مستوى الأداء لتقدير الحفظ يوضع ويحدد بالنسبة لكل هدف أو لكل واجب معين. ويجب أن يوضع هذا المستوى في مصطلحات لسرعة الأداء (مثل، يقيس الزاوية المنفرجة لأقرب جزء من الدرجة)، أو للنسبة المئوية من الفقرات التي يجيبها صحيحة (مثل، يعرف ٨٠٪ من المصطلحات).

وتستخدم النسبة المئوية للدرجة في الحكم عما اذا كان المتعلم قد حصل الأهداف، وعليها يتم تقرير نتيجة المتعلم في الاختبارات المعتمدة على المحك.

ويوضح الجدول التالي (جدول ١١) نموذجا لتقرير واحد من المتعلمين استخدمت معه النسبة المئوية. ويبدو من الجدول أن مستوى الحفظ وضع في التقرير مساويا ٨٠٪ وأن هذا المتعلم قد حصل كل الأهداف ما عدا واحدا فقط. وعلى الرغم من كون مستويات الأداء هذه كيفية الا أنها تظهر أسسا مفيدة لتخطيط عمل علاجي.[1]

جدول (١١)

تقرير فردي لمتعلم، عن اختبار حفظ قائم على محك

النسبة المئوية = ٨٠٪	المتعلم:		
الهدف- عدد الفقرات	عدد الفقرات الصحيحة	النسبة المئوية	تحقيق الهدف
١- يعرف المصطلحات (٢٠)	١٨	٩٠	×
٢- يعرف الخطوات (٢٠)	١٨	٩٠	×
٣- يفهم الأسس والمبادىء (٢٠)	١٧	٨٥	×
٤- يطبق الأسس والمبادىء (٢٠)	١٤	٦٠	
٥- يفسر البيانات (٢٠)	١٦	٨٠	×

واذا فشل غالبية المتعلمين في تحقيق هدف ما، فمن الممكن اذن تخطيط برنامج تعليم علاجي لكل أفراد المجموعة، أما اذا فشل عدد قليل منهم، فيكون اذن العلاج الفردي أكثر فاعلية.

ويعتبر تحليل الفقرة تفصيل متطلب ومفيد لتحليل نتائج الاختبار، والجدول التالي (جدول ١٢) مفيد في تحديد والتعرف على أخطاء التعلم، فبنظرة فاحصة أمام ما سجل لكل متعلم يمكن وضع اليد على أخطائه، وبنظرة أخرى في كل عمود

---

(١) Gronlund, N. (1973) **Preparing Criterion - Referenced Tests for Classroom Instruction.** New York: Macmillan Publishing Co.

يستطيع المعلم أيضا تقدير مستوى استجابة متعلميه ككل في كل فقرة من فقرات الاختبار. والجدول (جدول ١٢) هو نموذج فقط لعدد قليل من المتعلمين في عينة قليلة من فقرات أحد الاختبارات. [١]

ويفيد مثل هذا الجدول كثيرا في اختبارات تحديد الشكل أو الوقائية، والتحليل المفصل يزود واضع الاختبار بالمعارف التي يحتاجها لوضع وصفة تصحيح وعلاج لكل متعلم أو لمجموعة المتعلمين.

وبنظرة سريعة الى الجدول نلاحظ أن الفقرة (١٢) الثانية عشرة، يمكن تقدير لماذا أجاب جميع المتعلمين عنها اجابة خاطئة؟.

جدول (١٢)

مسلسل المتعلم	رقم الفقرة											
	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
١	/	/	/	/	/	/	/	/	×	/	×	×
٢	/	/	/	/	/	/	/	/	×	×	×	×
٣	/	/	×	/	/	/	×	×	/	/	/	×
٤	/	/	/	/	/	/	/	×	/	/	×	-
٥	/	/	/	/	×	×	×	×	/	/	×	×

تفسير الاختبارات المعتمدة على المعايير

تصمم هذه الاختبارات للتعرف على أداء المتعلم في اختبار ما ومقارنته بالآخرين، وتفسير هذه الاختبارات يتعلق بتقدير المستوى المناسب لكل متعلم

(١) محمد رضا البغدادي (١٩٧٩) «وحدات منهج الاحياء للصف الاول الثانوي في المملكة العربية السعودية- تقويم محتوياتها وطرق تدريسها» الكويت: جمعية المعلمين الكويتية. (اسبوع التربية التاسع ٢٤-١٩٧٩/٣/٣٠).

داخل حجرة الدراسة.

وسوف نقتصر مناقشتنا في هذا الجزء على بعض طرق معالجة الدرجات الخام، ومنها:

أ- ترتيب الدرجات الخام ترتيبا بسيطا

ان أبسط الطرق لتقديم درجات خام لأحد الاختبارات الى المتعلمين في حجرة الدراسة هي وضع هذه الدرجات في قائمة على السبورة، ويتم اجراء ذلك بترتيب هذه الدرجات تنازليا تبدأ بالدرجة المرتفعة متدرجة تنازليا الى المنخفضة، وذلك مع ضرورة عمل عد تكراري لبيان عدد (ن) المتعلمين الذين حصلوا على كل درجة من هذه الدرجات.

ويوضح الجدول التالي (جدول ١٣) هذا الأسلوب، حيث يتبين منه توزيع درجات ثلاثين متعلما (ن =٣٠) في اختبار يتألف من (٤٠) أربعين فقرة اختبارية. وعندما ينظر المتعلم الى مثل هذه القائمة ويعرفته للدرجة التي حصل عليها يستطيع عندئذ تحديد وضعه وموقعه بين أفراد مجموعته، أي يتعرف على مستواه فيما بين المجموعة.

جدول (١٣)

توزيع تكراري لدرجات خام في اختبار موضوعي يتألف من أربعين فقرة

التكرارات		الدرجات الخام
١		٣٨
١		٣٧
صفر		٣٦
٢		٣٥
١		٣٤
٢		٣٣
٣		٣٢
٢		٣١
١		٣٠
٤	الوسيط	٢٩
٢		٢٨
٢		٢٧
٢		٢٦
٣		٢٥
١		٢٤
صفر		٢٣
١		٢٢
١		٢١
صفر		٢٠
١		١٩
ن = ٣٠ ر		٨٧٢

ب- المعالجة الاحصائية لتوصيف مجموعة من الدرجات

يكون من المتطلب- أمام بعض الأغراض- توصيف مجموعة من الدرجات في شكل موجز عن وضعها في ترتيب شامل، ويكون ذلك باجراء عمليات حسابية احصائية لكل من:

- المتوسط أو النزعة المركزية
- انتشار أو توزيع الدرجات

وتتعدد آراء رجال الاحصاء في استخدام المصطلح «متوسط» في توصيف درجات الاختبار وذلك لأن هناك عدداً من الأنواع المختلفة للمتوسط. والأنواع الشائعة هي:

١- الوسيط Median ، أو حساب المتوسط، والذي يقدر بترتيب الدرجات تبعاً للحجم ثم حساب النقطة الوسط لسلسلة الدرجات.

٢- المتوسط Mean ، أو المتوسط الحسابي، والذي يقدر بترتيب وجمع جميع الدرجات وقسمة هذا المجموع على عدد هذه الدرجات.

٣- المنوال Mode ، وهو الدرجة التي تحدث كثيراً، ويقدر بفحص تكرارات كل درجة.

ان انتشار أو توزيع مجموعة من الدرجات يمكن بيانه بأكثر من طريقة، وأهم طريقتين هما:

أ- المدى Range .

وهو الحد الفاصل بين الدرجات العليا والدرجات الدنيا.

ب- الانحراف المعياري Standard deviation .

وهو المعدل الذي تنحرف فيه مجموعة من الدرجات عن المتوسط.

ويفهم الانحراف المعياري ببساطة بملاحظة أهميته في الاستخدامات

الاحصائية وفي توصيف انتشار الدرجات لمجموعة ما وفي قياس معاملات الثبات والأخطاء الثابتة في القياس وثبات الدرجات.

وعند توصيف درجات الاختبار يكون استخدام المدى والوسيط أو المتوسط ذا أهمية بينما يستخدم الانحراف المعياري مع المتوسط فقط.

جـ- تقدير الوسيط والمدى

ان أبسط طرق وصف درجات الاختبار هي استخدام الوسيط والمدى[1] وتستخدم هذه المقاييس عادة عندما تكون المجموعة قليلة العدد ولا توجد حاجة لاحصاءات أكثر. أي عندما تستخدم فقط في المناقشات داخل حجرات الدراسة لمعاونة المتعلمين في تفهم مواقعهم بين أفراد المجموعة. والخطوة الأولى هي ترتيب درجات الاختبار كما هو موضح في الجدول السابق (جدول ١٣)، ويقدر عندئذ الوسيط بتحديد نقطة الوسط لمجموعة الدرجات.

ويمكن اجراء ذلك في عمود التكرارات بعد النقطة الوسط بين التكرار الخامس عشر والسادس عشر. وفي هذه الحالة (كما ورد في جدول ١٣) تكون النقطة الوسط (الوسيط) واقعة أمام الدرجة (٢٩) ويحسب المدى بعملية حسابية بسيطة وذلك بطرح الدرجة الدنيا من الدرجة العليا (ويضاف الواحد الصحيح أحياناً)، وعليه يكون المدى لعدد ثلاثين متعلماً هو ١٩ (٣٨- ١٩)، أي أنه يمكن وصف مجموعة الدرجات الخام الواردة في الجدول أنها ذات وسيط يساوي ٢٩، ومدى يساوي ١٩. وهذا الوصف قد يكون مناسباً فقط لتوصيف درجات عينة قليلة من الأفراد ويصلح فقط لأغراض المناقشة داخل حجرات الدراسة، ولكنه لا يصلح للعمليات أو المعالجات الاحصائية[2].

---

(١) يمكن استخدام الانحراف الرباعي Quartil deviation بدلاً من المدى، ولكن يحدث ذلك نادراً في توصيف درجات اختبار المتعلمين داخل حجرات الدراسة.

(٢) السيد محمد خيري (١٩٧٠) مرجع سابق.

رمزية الغريب (١٩٧٠) مرجع سابق.

د- حساب المتوسط والانحراف المعياري

يقوم المتوسط والانحراف المعياري على قيمة كل درجة في المجموعة ، وهذا يعطي مقاييس ثابتة عن الأخرى التي نحصل عليها بالحساب، بالاضافة الى أن هذه المعالجات الاحصائية تستخدم في حساب عدد من الاحصاءات الأخرى ذات الفائدة في القياس والاختبار مثل مقاييس الثابت وغير ذلك. ولهذه الأسباب يفضل المتوسط والانحراف المعياري عند تحليل وتوصيف درجات الاختبار. ويمكن الحصول على المتوسط بعملية جمع حسابي لكل درجات المجموعة ثم قسمة المجموع على عدد الدرجات، ويمكن حساب المجموع (مجموع الدرجات) التي حصل عليها ثلاثون متعلمًا (جدول ١٢) فنجدها = ٨٧٢، وعند قسمة هذا المجموع على ٣٠ متعلمًا يكون المتوسط مساوياً ٢٩,٠٧ . وأبسط طرق حساب الانحراف المعياري تتضمن طرح مجموع أقل خمس درجات من مجموع أعلى خمس درجات، وقسمة الناتج على نصف عدد أفراد العينة، كما في المعادلة:

$$\text{الانحراف المعياري} = \frac{\text{مجموع أعلى خمس درجات} - \text{مجموع أقل خمس درجات}}{\text{نصف عدد أفراد العينة}}$$

$$\text{أي أن الانحراف المعياري} = \frac{١٨٠ - ١٠٥}{١٥} = ٥$$

وعليه يكون توصيف مجموعة الدرجات التي وردت في الجدول السابق انها ذات متوسط يدور حول الرقم (٢٩) وانحراف معياري (٥).

وعندما يكون هناك توزيع معتدل فمن المتوقع أن يكون ثلثا الدرجات تقع بين ٢٤ (٢٩-٥)، ٣٤ (٢٩ + ٥) أي أن حوالي ٧٠٪ من عدد الدرجات تقع خلال هذا المدى ، وعليه فان الانحراف المعياري يمكن استخدامه كدلالة تصف انتشار الدرجات في المجموعة.

---

(٣٦) Diederich, P.B. (1973) **Short - Cut Statistics for Teacher Made Tests.** Prince ton, N. Educational Testing Service.

كذلك يمكن الاستفادة من معامل الصعوبة ومعامل السهولة في حساب الانحراف المعياري لكل فقرة من فقرات الاختبار، ويساوي الانحراف المعياري في هذه الحالة:

$$\text{الانحراف المعياري} = \sqrt{\text{معامل السهولة} \times \text{معامل الصعوبة}}$$

هـ ـ استخدام النظام التساعي للدرجات المعيارية

كثيراً ما نميل أثناء تعليمنا الى مقارنة التحصيل النسبي للمتعلم في أكثر من اختبار، أو مقارنة مستواه في اختبار ما بمستواه في مقياس تحصيلي آخر. وأمام هذا الأمر نحتاج الى عناصر أخرى مثل الدرجات والترتيب.

والنظام التساعي ليس مفيداً للاختبارات التحصيلية فحسب، ولكنه يستخدم أيضاً مع الاختبارات المعيارية، حيث يمكن مقارنة درجات تحصيل المتعلم الراهنة مع هذه الاختبارات المعيارية من حيث التحصيل والاستعدادات.

والمقياس التساعي نظام للدرجات المعيارية يقسم توزيع الدرجات الخام الى أقسام تسع (ومن هنا كانت تسمية النظام)، أعلاها تسع الدرجة (٩)، وأقلها تسع الدرجة (١) وتسع الدرجة (٥) تقع في مركز التوزيع.

وتتوزع الدرجات المعيارية على التساعيات بمتوسط قدره (٥) وانحراف معياري قدره (٢)، ويمكن استخدام الجدول التالي (جدول ١٤) لتقدير عدد الأفراد في المجموعة الذين حصلوا أو الذين يحصلون على كل تساعي.

ولاستخدام هذا الجدول، كل ما نحتاجه هو الدخول الى الجدول بعدد أفراد المجموعة (حجم المجموعة) وقراءة الجدول على طول الخط، مع ملاحظة عدد الأفراد الذين يحصلون على كل تساعي، فمثلاً، اذا كان حجم المجموعة (٣٠) ثلاثين متعلمًا، يحصل متعلم واحد على درجة تساعية (١)، واثنين من المتعلمين على درجة تساعية (٢)، وأربعة متعلمين على درجة تساعية (٣)، وهكذا. وعندما تتحول الدرجات الخام الى درجات تساعية تصبح في موضع فعّال على المقياس

المعياري، وعليه فان الفرق بين التساعي (٧) والتساعي (٨) هو نفس الفرق بين التساعي (٤) والتساعي (٥)، يلاحظ في الجدول التالي (جدول ١٣) ان حجم العينة هو ٦٠ ستون متعلمًا، فاذا كانت العينة التي تعمل معها أكثر من هذا العدد فارجع الى «جرونلند Gronlund »[1] .

(١) Gronlund, N.F. (1976) **Measurment and Evaluation in Teaching.** مرجع سابق.

obeikandl.com

جدول (١٤)

عدد الأفراد وحصولهم على الدرجات التساعية

حجم المجموعة	التساعيات								
	١	٢	٣	٤	٥	٦	٧	٨	٩
	عدد الأفراد الذين يحصلون على الدرجات التساعية								
٢٠	١	١	٢	٤	٤	٤	٢	١	١
٢١	١	١	٢	٤	٥	٤	٢	١	١
٢٢	١	٢	٢	٤	٤	٤	٢	٢	١
٢٣	١	٢	٢	٤	٥	٤	٢	٢	١
٢٤	١	٢	٣	٤	٤	٤	٢	٢	١
٢٥	١	٢	٣	٤	٥	٤	٣	٢	١
٢٦	١	٢	٣	٤	٦	٤	٣	٢	١
٢٧	١	٢	٣	٥	٥	٥	٣	٢	١
٢٨	١	٢	٣	٥	٦	٥	٣	٢	١
٢٩	١	٢	٤	٥	٥	٥	٣	٢	١
٣٠	١	٢	٤	٥	٦	٥	٤	٢	١
٣١	١	٢	٤	٥	٧	٥	٤	٢	١
٣٢	١	٢	٤	٦	٦	٦	٤	٢	١
٣٣	١	٢	٤	٦	٧	٦	٤	٢	١
٣٤	١	٣	٤	٦	٦	٦	٤	٣	١
٣٥	١	٣	٤	٦	٧	٦	٤	٣	١
٣٦	١	٣	٤	٦	٨	٦	٤	٣	١
٣٧	٢	٣	٤	٦	٧	٦	٤	٣	٢
٣٨	١	٣	٥	٦	٨	٦	٤	٣	١
٣٩	١	٣	٥	٧	٧	٧	٥	٣	١
٤٠	١	٣	٥	٧	٨	٧	٥	٣	١
٤١	١	٣	٥	٧	٩	٧	٥	٣	١
٤٢	٢	٣	٥	٧	٨	٧	٥	٣	٢

٤٣	٢	٣	٥	٧	٩	٧	٥	٣	٢
٤٤	٢	٣	٥	٨	٨	٨	٥	٣	٢
٤٥	٢	٣	٥	٨	٩	٨	٥	٣	٢
٤٦	٢	٣	٥	٨	١٠	٨	٥	٣	٢
٤٧	٢	٣	٦	٨	٩	٨	٥	٣	٢
٤٨	٢	٣	٦	٨	١٠	٨	٥	٣	٢
٤٩	٢	٤	٦	٨	٩	٨	٦	٤	٢
٥٠	٢	٣	٦	٩	١٠	٩	٦	٣	٢
٥١	٢	٣	٦	٩	١١	٩	٦	٣	٢
٥٢	٢	٤	٦	٩	١٠	٩	٦	٤	٢
٥٣	٢	٤	٦	٩	١١	٩	٦	٤	٢
٥٤	٢	٤	٧	٩	١٠	٩	٦	٤	٢
٥٥	٢	٤	٧	٩	١١	٩	٧	٤	٢
٥٦	٢	٤	٧	٩	١٢	٩	٧	٤	٢
٥٧	٢	٤	٧	٩	١١	١٠	٧	٤	٢
٥٨	٢	٤	٧	١٠	١٢	١٠	٧	٤	٢
٥٩	٣	٤	٧	١٠	١١	١٠	٧	٤	٣
٦٠	٣	٤	٧	١٠	١٢	١٠	٧	٤	٣

الصدق والثبات

يعزى الصدق الى ملاءمة وموافقة ترجمة وتفسير نتائج الاختبار، ويعزى الثبات الى تماسك واتساق هذه النتائج، وكلا الصدق والثبات مفاهيم هامة في تحسين واستخدام الاختبارات التحصيلية.

فالصدق والثبات خاصتان ضروريتان وجديرتان بالاعتبار عند إعداد واستخدام اختبار التحصيل، فهما اذن ضرورتان للقياس الفعال يجب أن يلم بها كل مشتغل في الاختبارات. وعليه فإن فهم وادراك هذين المفهومين ينمي المهارة المتطلبة في تصميم الاختبار وتفسير نتائجه.

تقدير الصدق

يتصل الصدق بمدى ما اذا كانت نتائج الاختبار تخدم الغرض والاستخدام المتطلب. فمثلا، يجب أن تستخدم نتائج الاختبار في توصيف أنواع المعارف التي يمكن أن يظهرها المتعلم الفرد، أو للتخطيط لنشاطات جديدة.

ويتصل سؤ النا بماهية الاستجابات السلوكية التي يقيسها الاختبار، وبمدى دقة الدرجات وارتباطها بأنواع الأداء الأخرى ذات القيمة. وعليه فإن نتائج الاختبار ذات فوائد واستخدامات عدة، وهناك ألوان مختلفة من الاستدلالات تؤخذ عن نتائج هذه الاختبارات، ولذلك يعزى الصدق الى الملاءمة والموافقة في ترجمة وتفسير هذه النتائج.

ومفهوم الصدق كما هو مستخدم في القياس والاختبار يمكن توضيحه بالاشارة الى النقاط التالية:

١- يعزى الصدق الى ترجمة وتفسير نتائج الاختبار وليس للاختبار نفسه.
٢- يستدل على الصدق من الدلائل والعلامات المتوفرة والتي لا تقاس.
٣- للصدق فاعلية في استخدامات مثل الاختيار والتقويم وتحديد الوضع.
٤- يعبر عن الصدق بدرجات لفظية مثل عال ومتوسط ومنخفض.

وقد قسمت الجمعية النفسية الأمريكية [١] الصدق الى الأقسام الرئيسة التالية:

١- صدق المحتوى

٢- الصدق المرتبط بالمحك (التنبؤي- التلازمي)

٣- صدق التصميم

---

(١) American Psychological Association (1974) **Standards for Educational and Psychological Tests.** Washington, D. C.:
انظر الفصل الخاص بالصدق، ص ٢٥
، الفصل الخاص بالثبات، ص ٤٨

فصدق المحتوى يتناول دراسة مفردات الاختبار ومحتوياته ومادته، والاختبار الصادق اذن هو الاختبار الذي يمثل تمثيلا سليما الميادين والمجالات المراد قياسها ودراستها والصدق المرتبط بالمحك لونان، الأول عن كيفية التنبؤ بالأداء في الاختبار (صدق تنبؤي) وتقدير المستوى الراهن بالمقارنة باختبار أو مقياس مرجعي أو محك من أي نوع من الأنواع المعترف بها (صدق التلازم).

أما صدق التصميم، فهو كيفية توضيح الأداء في اختبار ما بمصطلحات ذات صفات وخواص نفسية.

صدق المحتوى

نهتم كثيرا بهذا النوع من الصدق عندما نريد استخدام الأداء في الاختبار كدليل على تمثيل عينة ممثلة من المواقف. فمثلا، اذا كانت لدينا قائمة من ٥٠٠ كلمة نتوقع من المتعلمين القدرة على هجائها صحيحا عند نهاية العام الدراسي، فلقياس قدراتهم على الهجاء يمكن اعطاؤهم اختبارا هجائيا يتألف من خمسين كلمة تمثل النوعيات المختلفة لكلمات القائمة. وسوف يكون أداؤهم في هذه الكلمات ذا أهمية كعينة فقط وكعلامة ودليل على قدراتهم على هجاء القائمة جميعها. وعليه فإن اختبار الهجاء هذا يجب أن يكون له صدق محتوى. فاذا تخيرنا الكلمات السهلة فقط أو الكلمات الصعبة فقط أو الكلمات ذات الأخطاء الشائعة فان هذا الاختبار يحتاج الى صدق محتوى قليل، أما اذا تخيرنا عينة متوازنة من الكلمات فسوف يحتاج الاختبار الى صدق محتوى عال.

يتضح من العرض السابق أن العنصر الفعال في صدق المحتوى هو الكفاية في العينة. فالاختبار عادة عينة من العديد من الأسئلة التي يجب أن توجه، أي أن هذا الصدق هو الحالة في تقدير ما اذا كانت العينة ممثلة للعينة الكلية أم لا.

ويعتبر صدق المحتوى ذا أهمية في الاختبارات التحصيلية، أي في تمثيل فقرات الاختبار للموضوعات والمعطيات التعليمية التي تمت تغطيتها خلال الفترة التعليمية. وعليه فيمكن بناء وتصميم اختبار على درجة عالية (طيبة) من صدق المحتوى وذلك عند:

١ـ تحديد وتعيين الموضوعات الرئيسة الدراسية والمعطيات السلوكية التي ستخضع للقياس.

٢ـ تصميم وإعداد جدول مواصفات يعين فيه نوعية الفقرات.

٣ـ مطابقة فقرات الاختبار بقائمة المواصفات.

فهي خطوات تضمن الى حد بعيد قياس العينة الممثلة لكل من المادة التعليمية والمعطيات السلوكية ودرجة عالية من صدق المحتوى في نفس الوقت.

يتبع واضعو الاختبارات المقننة أو المعيارية هذه الخطوات في بناء وتصميم الاختبارات التحصيلية، ولكن المحتوى والمعطيات السلوكية المتضمنة في قائمة المواصفات (جدول المواصفات) تقوم على مدى أكثر اتساعا وانتشارا. فهي تقوم على الكتاب المدرسي المقرر وعلى توصيات وآراء الخبراء والمتخصصين في المجالات التي يغطيها الاختبار. وعليه فالاختبار التحصيلي المقنن يجب أن يكون على درجة عالية من صدق المحتوى بصفة عامة، فيجب أن يتألف من المحتويات التعليمية المنتشرة في مجال الدراسة.

الصدق المرتبط بالمحك

هناك نوعان لهذا الصدق، الأول يتصل كثيرا باستخدام أداء الاختبار للتنبؤ بأداء لاحق على مقاييس معترف بها تعرف بالمحكات. فمثلا، نستخدم درجات اختبار الاستعداد المدرسي للتنبؤ بالمستوى الدراسي كمحك، ولذلك يسمى هذا النوع بالصدق التنبؤي.

أما النوع الثاني فيتصل باستخدام أداء الاختبار لتقدير أو لحساب الأداء الشائع في بعض المحكات، فمثلا عندما نريد استخدام اختبار لدراسة المهارات لتقدير وحساب ما هو المعطى الذي يكون موضع ملاحظة دقيقة من المتعلمين في دراسة واقعية كمحك.

ومن هنا يمكن القول أن الحصول على كلا النوعين يكون في وقت واحد ولذلك

سمي هذا النوع بصدق التلازم.

ويمكن القول بصفة عامة أن القاسم المشترك في كلا النوعين من الصدق المرتبط بالمحك هو درجة العلاقة بين مجموعتي القياس (درجات الاختبار التحصيلي والمعيار)، هذه العلاقة هي التي تعرف بمعامل الارتباط أو بقائمة الاحتمالات.

معاملات الارتباط

يعين معامل الارتباط درجة العلاقة بين اختبارين أو بين مجموعتين من المقاييس، ويظهر الارتباط الموجب عندما تكون درجات أحد الاختبارات عالية ومصاحبة لدرجات عالية للاختبار الآخر، ويظهر الارتباط السالب عندما تكون درجات الاختبارين منخفضة، أي أن

ارتباط موجب (علاقة ايجابية) = +١,

ارتباط سالب (علاقة سالبة) = -١,

لا ارتباط (لا علاقة) = ٠٠,٠٠

عندما يستخدم معامل الارتباط لتوضيح درجة العلاقة بين مجموعتين من درجات اختبار ومقياس معياري فيسمى هذا المعامل بمعامل الصدق. فمثلا، معامل صدق يساوي (+١,) يعني أن العلاقة بين مجموعة درجات اختبار استعداد (المتنبىء) ومجموعة درجات اختبار تحصيل (المعيار) تشير الى أن كل متعلم في المجموعة له نفس مستوى الصلة لكلا الاختبارين، وبذلك يقدم معامل الصدق تنبؤ أصادقاً صحيحاً من درجات الاستعداد الى درجات التحصيل. والجدير بالاشارة الى أن معظم معاملات الصدق أقل من الواحد الصحيح، وعلى أية حال فإن المدى الموجب يعطي علامة هداية مفيدة لتقويم معامل الصدق.

وتتلخص الخطوات التي تستخدم لتقويم معامل الصدق في مقارنته بمعاملات الصدق التي يحصل عليها عندما يكون هناك مقياسان مرتبطان. فمثلا، معامل صدق (٤٠,) بين درجات اختبار استعداد ودرجات اختبار تحصيل يعتبر معاملا

صغيرا، لأننا عادة نحصل على ارتباطات تدور بين (٥، ٧,) بين هذين المقياسين. وعليه يمكن القول أن استخدام معاملات الصدق بفاعلية يتم عندما يكون أحدهما على صلة بقيمة معاملات الصدق التي يحصل عليها بين العديد من أزواج المقاييس.

قائمة ـ جدول ـ الاحتمالات

تعد قائمة الاحتمالات ذات فائدة لعمل تنبؤات من درجات الاختبار، وهي مخطط ثنائي من درجات الاختبار (المتنبىء) مرتبة أسفل الجانب الأيمن، ودرجات المعيار مرتبة في مستويات، كل مستوى عبارة عن النسبة المئوية من المتعلمين تبعا لكل مستوى في المعيار. ونعرض المثال التالي، كما هو موضح في الجدول التالي (جدول ١٥)

جدول (١٥)

قائمة (جدول) الاحتمالات، توضح العلاقة بين درجات الاستعداد المدرسي ودرجات الاختبار لثلاثين متعلما في مقرر للعلوم

درجات مجموعات الاستعداد المدرسي (تساعيات)	النسب المئوية في كل مستوى تبعا لكل معيار				
	أ	ب	ج	د	هـ
فوق المتوسط (٧، ٨، ٩)	٤٣	٤٣	١٤		
متوسط (٤، ٥، ٦)	١٩	٢٥	٣٧	١٩	
تحت المتوسط (١، ٢، ٣)			١٤	٢٩	٥٧

يلاحظ أن هؤلاء المتعلمين الذين يقعون في مجموعة (فوق المتوسط، تساعي ٧، ٨، ٩) في درجات الاختبار حصل ٤٣٪ منهم على المستوى «أ»، ٤٣٪ على المستوى «ب»، و ١٤٪ على المستوى «جـ»، ولما كانت هذه النسب تقوم على أساس هذه المجموعة فإنه من الممكن استخدامها للتنبؤ بالأداء اللاحق للمجموعات الأخرى من المتعلمين في مقرر العلوم. واذا وقع متعلم ما في مجموعة (فوق المتوسط) في اختبار الاستعداد المدرسي فسوف نتنبأ له بأن يكون أمامه ٤٣ فرصة من ١٥٠

فرصة ليستأهل المستوى (أ)، ٤٣ فرصة من ١٠٠ فرصة ليستأهل المستوى (ب)، ١٤ فرصة من ١٠٠ فرصة ليستأهل المستوى (ج) في هذا المقرر المعين.

وعليه، يبدو أن قائمة الاحتمالات سهلة الإعداد سريعة الاستخدام في بيان العلاقة بين مجموعات الدرجات.

طرق أخرى غير مباشرة لحساب معامل الصدق

اذا لم يوجد معيار أو محك له درجة عالية من الثبات، استخدمت طرق اخرى غير مباشرة في حساب معامل الصدق، على أن هذه الطرق أقل دقة من الطرق الثابتة (بواسطة معاملات الارتباط أو بواسطة جدول الاحتمالات) السابقة، وتستخدم حين يراد حساب معامل صدق تقريبي، ومن هذه الطرق:

١ـ تعطى مجموعة من الاختبارات يفترض فيها أنها تقيس نفس السمة أو نفس الظاهرة لعينة ممثلة من الأفراد، ثم يحسب متوسط ارتباط كل اختبار بالاختبارات الأخرى، ويعتبر أكبرها ارتباطا أكثرها صدقا.

٢ـ تعطى الاختبارات التي تقيس سمة أو قدرة معينة لعينة ممثلة ثم تجمع درجات الاختبارات جميعا، ويحسب معامل صدق كل اختبار بواسطة معامل ارتباطه مع هذه الدرجات مجتمعة.

٣ـ طريقة أخرى، وهي الاعتماد على عدد من الحكام أو الخبراء المعروف بكفايتهم، وتستخدم هذه بكثرة في تصميم استفتاءات الشخصية المتعلقة بالسمات العصابية المختلفة.

٤ـ استخدام ما يعرف بالصدق الذاتي، وهو الذي يعطينا معامل الارتباط بين الدرجات التجريبية أو درجات الاختبار الجاري تقنينه وبين الدرجات الحقيقية، فالصدق الذاتي هو في الواقع معامل صدق لا بأس به، لأنه يدلنا على مدى نجاح الاختبار في قياس القدرة الحقيقية.

والجدير بالذكر، انه هناك بعض العوامل التي يمكن أن تؤثر على صدق

الاختبار، وبالتالي على قدرته على التنبؤ بالتحصيل في المعيار أو المحك. ومن هذه العوامل طول الاختبار، وثبات الاختبار، وثبات المرجع، والتباين واقتران ثبات الميزان بثبات الاختبار.

تقدير الثبات

يعزى الثبات الى اتساق وتجانس درجات الاختبار، أي كيف تتسق درجات الاختبار؟ ويمكن توقع تغيرات عديدة في أداء الاختبار من وقت لآخر، ومن عينة من الفقرات الى أخرى ومن جزء من الاختبار الى جزء آخر. وقياسات الثبات تزود بتقدير مدى الانحراف الذي يمكن توقعه تحت الظروف المختلفة. ويقرر ثبات درجات الاختبار بما يعرف بمعامل الثبات أو الخطأ المعياري للقياس الذي يشتق منه.

وكما أسلفنا الذكر بأن معامل الارتباط يوضح العلاقة بين مجموعة درجات اختبار ومقياس معياري وهو معامل الصدق، فإن معامل الثبات أيضا معامل ارتباط ولكن يظهر الارتباط بين مجموعتين من المقاييس تأخذ نفس الأسلوب ونفس الخطوات.

وقد يكون ذلك بتطبيق نفس الاختبار مرتين على أفراد نفس المجموعة من المتعلمين مع وجود فترة زمنية فاصلة (وتعرف بطريقة اعادة تطبيق الاختبار)، أو بتطبيق صورتين متكافئتين من الاختبار تطبيقا متتاليا (وتعرف بطريقة الصور المتكافئة)، أو بتطبيق صورتين متكافئتين من الاختبار مع وجود فترة زمنية فاصلة (وتعرف بطريقة اعادة تطبيق صورتين متكافئتين من الاختبار)، أو بتطبيق الاختبار مرة واحدة ويحسب الاتساق من داخل الاختبار نفسه (وتعرف بطريقة الاتساق الداخلي). ومن كل طريقة من هذه الطرق نحصل على معامل ثبات بشكل يخدم نوعية معينة من المعارف. وهناك أربعة طرق أساسية شائعة الاستخدام لحساب معامل الثبات هي:

١ـ طريقة اعادة تطبيق الاختبار

وتعتمد هذه الطريقة على اجراء الاختبار على عينة كافية وممثلة من المتعلمين،

ثم اعادة تطبيقه مرة أخري عليهم، بشرط أن يمر وقت مناسب لا يؤثر على تحصيل أفراد العينة في المرة الثانية بالألفة والتمرين.

وبعد اجراء التطبيق الثاني، يحسب معامل الارتباط بين الأداء في المرتين، فيكون معامل الارتباط هو معامل الثبات، فاذا كان عالياً دل ذلك على أن الأداء في المرة الثانية لم يكن مختلفا عن الأداء في المرة الأولى الا بما يسمح به التأثر بأخطاء الصدفة، وان كان صغيرا دل ذلك على اختلاف الدرجات في التحصيل في المرتين، وكان الاختبار غير ثابت، ولا يمكن الاعتماد عليه.

وتعتبر هذه الطريقة من أفضل الطرق المستخدمة في حساب معامل ثبات الاختبارات غير الموقوتة.

ومن أهم ما يوجه الى هذه الطريقة من مآخذ هو أن استجابات المتعلمين تتأثر بعامل الألفة والتمرين على الاختبار، ولذلك كانت الدرجات في المرة الثانية بصفة عامة أعلى منها في المرة الأولى. ولذلك لا تستخدم هذه الطريقة في قياس معاملات الثبات مع اختبارات التذكر وما شابهها لأن اعادة الاختبار نفسه مرة أخرى يؤثر في قياس هذه القدرة.

ولذلك فإننا نتفق مع الآراء التي ترى أن الأداء في كل مرة يعاد فيها الاختبار يتأثر ببعض العوامل مثل عامل التعب والصحة والدافعية، وغير ذلك. اذ لا يمكن التحكم تماما في هذه الظروف في حالتي اجراء الاختبار بحيث تكون المرة الثانية صورة مطابقة تماما للظروف في المرة الأولى، لهذا كانت النتائج التي نحصل عليها باستخدام هذه الطريقة مشوبة ببعض العيوب التي تجعلها غير دقيقة.

٢ـ طريقة الصور المتكافئة

وتستخدم هذه الطريقة حين يتوفر عند مصمم الاختبارات، اختبارات متكافئة تماما، بحيث تتفق في القيم المختلفة لمقاييسها الاحصائية مثل المتوسط والانحراف المعياري ومعامل الارتباط ومعامل الصدق، وتقوم فكرتها على نفس

فكرة التجزئة النصفية التي يراد بها تحقيق التكافؤ بقدر الامكان [1]. وهذه الطريقة مثل طريقة التجزئة النصفية لا تصلح لقياس الاختبارات الموقوتة التي تأخذ عامل السرعة بعين الاعتبار.

٣ـ طريقة اعادة تطبيق الاختبار مع الصور المتكافئة

وهي طريقة ندمج فيها بين الطريقتين الأوليين، حيث يكون لدينا صورتين مختلفتين من نفس الاختبار تقدما على التتالي بفاصل زمني بينهما. وهذه الطريقة تفضل كثيرا من الطرق حيث أنها تضع في الاعتبار كل مصادر التغير الممكنة والتي تؤثر على التطبيق وبالتالي على المعاملات. ولذلك فهي تعتبر أكثر الأنواع لغالبية الأغراض. ومعامل الثبات العالي الذي يحصل عليه من هذه الطريقة يدل على أن درجات الاختبار لا تمثل فقط الأداء في الاختبار الراهن ولكن الأداء في الاختبارات الأخرى التي تتماثل معه في وقت آخر أو على عينة مختلفة من الفقرات المتكافئة.

٤ـ طريقة الاتساق الداخلي

وتتطلب هذه الطريقة مرة واحدة لعرض الاختبار، أي خطوة واحدة فقط. وتعتبر طريقة التقسيم للاختبار الى مجموعتين أو نما تعرف به طريقة التجزئة النصفية للاختبار أكثر الطرق استخداما لحساب ثبات الاختبار. ويرجع السبب الى أنها تتلافى عيوب بعض الطرق الأخرى، فهي تتلافى ما يوجه الى طريقة اعادة الاختبار من عيوب والتي أهمها أننا باعادة تطبيق الاختبار لا نضمن أن تكون ظروف التطبيقين واحدة، هذا فضلا عن أن اعادة تطبيق الاختبار قد يؤدي الى ألفة المتعلمين بالاختبار في التطبيق الثاني من التطبيق الأول. فاذا أضفنا الى ذلك كثرة تكاليفها وطول الوقت المستخدم في الاختبار لأمكننا أن نتبين لماذا كانت طريقة التجزئة النصفية مفضلة عنها.

كذلك تفضل هذه الطريقة على طريقتي الصورتين المتكافئتين لأنها أرخص

---

(١) نعيم عطية (١٩٧٠) التقييم التربوي الهادف. بيروت: دار الكتاب اللبناني.

وأسرع، وكل ما يؤخذ عليها أنها لا تسمح باختبار الفرد الا مرة واحدة، فاذا تصادف وجود ظروف معرقلة للتحصيل في الاختبار تأثرت بها النتائج.

وقد استخدمت قوانين مختلفة لحساب ثبات الاختبار بطريقة التقسيم النصفي منها:[1]

١ـ قانون سبيرمان براون

$$\checkmark_{11} = \frac{2 \times ر_{21}}{1 + ر_{21}}$$

٢ـ قانون فلانجان

$$\checkmark_{11} = \frac{4 ع_1 \times ع_2 \times ر_{21}}{ع_1^2 + ع_2^2 + 2 ع_1 ع_2 ر_{21}}$$

٣ـ قانون رولون

$$\checkmark_{11} = 1 - \frac{ع_3^2 (1 - 2)}{ع^2}$$

٤ـ قانون موسير

$$\checkmark_{11} = \frac{4 (ع_1 ع_2 ر_{12} - ع_1^2)}{ع^2 ن}$$

٥ـ قانون جتمان

$$\checkmark_{11} = 2 \left(1 - \frac{ع_1^2 + ع_2^2}{ع^2}\right)$$

ومن أبسط طرق حساب معامل الثبات للاختبار الذي يطبق مرة واحدة هي استخدام معادلة «كيودر ردتشاردسن رقم ٢١

(٢) Kuder- Richardson Formula 21

---

(١) رمزية الغريب (١٩٧٠). مرجع سابق.

(٢) السيد محمد خيري (١٩٧٠) م مرجع سابق.

obeikandl.com

والمعادلة هي:

$$KR\ 21 = I - \frac{M\ (K - M)}{K(S)^2}$$

ويمكن ترجمتها الى:

معامل الثبات (بطريقة كيودر ـ ٢١)

= ١ـ المتوسط الحسابي (عدد فقرات الاختبار - المتوسط الحسابي)

عدد فقرات الاختبار × مربع الانحراف المعياري

وتعطينا معادلة «كيودر» تقديرا معتدلا للثبات، حيث تقوم على الاتساق الذي يتم من استجابة المتعلم من فقرة الى أخرى في الاختبار.

وتستخدم بصفة عامة طريقة الاتساق الداخلي في حساب معامل الثبات مع الاختبارات التي لا تتطلب السرعة لأنه يحصل منها مع هذه النوعية من الاختبارات على معامل ثبات زائف.

* * * * * *

تم بحمد الله

قائمة المراجع

أولا: المراجع العربية

١ـ ابراهيم بسيوني عميرة وفتحي الديب (١٩٧٦) تدريس العلوم والتربية العلمية. القاهرة: دار المعارف بمصر.

٢ـ السيد محمد خيري (١٩٧٠) الاحصاء في البحوث النفسية والتربوية والاجتماعية. القاهرة: دار النهضة العربية.

٣ـ رمزية الغريب (١٩٧٠) التقويم والقياس النفسي والتربوي. القاهرة: مكتبة الأنجلو المصرية.

٤ـ محمد رضا البغدادي (١٩٧٦) «مقارنة مدى فاعلية عدة طرق في تدريس وحدة الوراثة لطلاب دور المعلمين والمعلمات». رسالة دكتوراه. جمهورية مصر العربية: جامعة أسيوط، كلية التربية بالمنيا.

٥ـ محمد رضا البغدادي (١٣٩٧ هـ) التعليم المبرمج. المملكة العربية السعودية: جامعة الرياض.

٦ـ محمد رضا البغدادي (١٩٧٩) «وحدات منهج الاحياء للصف الأول الثانوي في المملكة العربية السعودية، تقويم محتوياتها وطرق تدريسها». الكويت: جمعية المعلمين الكويتية. (اسبوع التربية التاسع ٢٤ـ ٣٠/٣/ ١٩٧٩).

٧ـ محمد رضا البغدادي (١٩٧٩) التدريس المصغر ـ برنامج لتعليم مهارات التدريس. الكويت: مكتبة الفلاح.

٨ـ نعيم عطية (١٩٧٠) التقييم التربوي الهادف. بيروت دار الكتاب اللبناني.

ثانياً: المراجع الأجنبية

9- American Psychological Association (1974) **Standards for Educational and Pschological Tests.** Washington, D. D.

10- Baldwin. T. S. (1971) «A **Evaluation of Learining in Industrial Education.** New York: McGraw Hill Book Co.

11- Bloom, B. S. (1965) **Taxonomy of Educational Objectives,** Handbook I, Cognitive Domain. New York: David McKay Co., Inc.

12- Bloom, B. S., J. T. Hastings, and G. F. Madaus (1971) **Handbook on formative and summative Evalutation of Student Learning.** New York: McGraw-Hill book Co.

13- Coffman, W. E. (1971) «Essay Examination» In: **Educational Measurment.** ed., R. I., Thorndike, Chap. 10. Washington, D. C. American Concil of Education.

14- Diederich. P. B. (1973) **Short-Cut Statistics for Teacher Made Tests.** Princeton, N. J.: Educational Testing Service.

15- Encyclopedia of Educational Research. London: The Macmillan Company.

16- Godshalk, F. I., F. Swwineneford and W. E. Coffman (1966) **The Measnrement of Writing Ability:** New York: College Entrance Examination Board.

17- Gronlund, N. (1970) **Stating Behavioral Objectives for Classroom Instruction.** New York: Macmillan Publishing Co., Inc.

18- Gronlund, N. (1973) **Preparing Tests for Classroom Instruction.** New York: Macmillan Co., Inc.

19- Gronlund, N. (176) **Measurement and Evaluation in Teaching.** (3 rd. ed.) New York: Macmillan Co., Inc.

20- Harrow, A. (1972) **Taxonomy of Psychomotor Domain.** New York: David McKay Co., Inc.

21- Krathwohl, D., et al., (1964) **Taxonomy of Educational objectives.,** Affective Domain. New York: David McKay Co., Inc.

22- Kryspin, W. J. and J. T. Feldhusen (1974) **Developing Classroom Tests.** Menneapolis: Burgess Publishing Co.

23- Mores, H. T., and G. M. McGune (1971) **Selected Items for the Testing-of Study Skills and Critical Thinking.** Washington, D. C., National Council for the Social Studies.

24- Simpson, E? J. (1972) -The Classification of Educational Objectives in the Psychomotor Domain» Vol. 3. Washington: Gryphon House.

25- Story. A. (1975) **Measurement Of Learning.** U. S. A.: Library of Congress.

26- Throndike, R. L., ed., (1971) **Educational Measurment.** Washington: American Council on Education.

27- Traxler, A. (1930) «Administering and Scoring the Objective Tests» In: **Educational Measurment,** ed., by Lindquisi. Vol. 32. PP, 142-157.

28- Wasman. A. G. (1971) Writing the Test Item In: **Educational Measurment** (2 nd, ed.,) ed. R. L. Thorndike., Washington: American Council Education.

رقم الإيداع	١٩٨٣/٤٠٣٦
الترقيم الدولى ٥ - ٠٥٧٤ - ٠٢ - ISBN ٩٧٧	

٣/٨٣/٣٧

طبع بمطابع دار المعارف (ج.م.ع.)

obeikandl.com