

تحديد خصائص الاختبار

Specifying the Characteristics of a test

أثناء هذه الفترة الانتقالية من تدريبات الاختبار ، تم تصميم وبناء العديد من الاختبارات باستخدام مبادئ النظرية التقليدية للاختبار ، ولكن تم تحليلها من خلال إجراءات نظرية الاستجابة للمفردة . وقد أدى عدم التوافق بين إجراءات وضع الاختبار وإجراءات التحليل إلى حماية نظرية الاستجابة للمفردة من أن تستغل حيث يجب تصميم الاختبارات ووضعها وتحليلها وتفسيرها في إطار النظرية . وبذلك فالهدف من هذا الفصل هو تزويد القارئ بالخبرة في الجوانب الفنية لوضع الاختبار ضمن إطار نظرية الاستجابة للمفردة .

ويقوم الأفراد الذين يتولون وضع الاختبار بعمل ذلك في العديد من المجالات المختلفة فيقومون بتطوير اختبارات للشركات التي تجري اختبارات ، وكذا المؤسسات الحكومية والمناطق التعليمية وبالإضافة إلى ذلك يقوم المعلمون في كل مستويات الصف الدراسي بتطوير اختبارات لقياس التحصيل . وفي كل هذه المواقف تستند عملية وضع الاختبار عادة إلى مجموعة من المفردات التي يتم اختيارها ، والتي تمثل في اختبار معين . وتعرف هذه المجموعات من المفردات بينوك المفردات . ويتم اختيار المفردات من هذه القوائم على أساس كل من محتواها وخصائصها الفنية أي القيم المعلمية لمفرداتها .

ووفقاً لنظرية الاستجابة للمفردة يتم استخدام مجموعة محددة من الإجراءات في إنشاء هذه القوائم والحفاظ عليها . وقد سميت هذه الإجراءات باسم خاص وهو بنوك الأسئلة . والهدف الأساسي هو الحصول على بنك أسئلة يتم فيه التعبير عن قيم معالم المفردات بمقياس قياس القدرة فإذا تم ذلك يمكن اختيار المفردات من بنك الأسئلة وتحديد الخصائص الفنية العامة للاختبار قبل إجرائه على مجموعة من الأفراد الذين يتم اختبارهم . وإذا لم تفي خصائص الاختبار بأهداف التصميم يمكن استبدال المفردات المختارة بمفردات أخرى من بنك الأسئلة حتى يتم الحصول على الخصائص المرغوبة . وبهذه الطريقة نوفر قسراً كبيراً من الوقت والمال يتم تخصيصه لإجراء الاختبار .

ولبناء بنك المفردات يلزم أولاً : تحديد السمة الكامنة التي تقيسها المفردات وكتابة المفردات لقياس هذه السمة وإجراء اختبار للمفردات لاستبعاد المفردات الضعيفة . وبعد فترة من الوقت تكون مجموعة المفردات التي تقيس السمة الكامنة متاحة ثم يتم تطبيق هذه المجموعة الكبيرة من المفردات على مجموعة كبيرة من الأشخاص الذين يتم اختبارهم . ويتم اختبار نموذج المنحنى المميز للمفردة وتحليل إجابة المفحوص من خلال نموذج بيرنبوم كما يتم معايرة الاختبار . ويعتبر مقياس القدرة الناتج عن هذه المعايرة هو المقياس الأساسي لبنك المفردات ومن وجهة نظر واضع الاختبار لدينا الآن مجموعة من المفردات نعلم القيمة المعلمية لها ، وهذا يعني من الناحية الفنية وجود " قائمة مفردات سبق معايرتها " .

تطوير اختبار من قائمة مفردات سبق معايرتها

Developing a Test From a Precalibrated Item Pool

وحيث إن المفردات في قائمة المفردات السابق معايرتها تقيس سمة كامنة خاصة ، فالاختبارات التي تُبنى ستقيس أيضاً هذه السمة . وبينما قد يبدو ذلك شاذاً

قليلاً إلا أننا نجد عدداً من الأسباب لوجود اختبارات إضافية تقوم بقياس نفس السمة . وعلى سبيل المثال نجد من الضروري وجود نماذج بديلة بطريقة معتادة لاستمرار سرية الاختبار وإمكانية استخدام النسخ الخاصة للاختبار للحصول على المنح الدراسية . وسيتم في هذه الحالات انتقاء المفردات من قائمة المفردات على أساس مضمونها وخصائصها الفنية لمواجهة أهداف الاختبار الخاصة . وتكمن ميزة وجود قائمة مفردات سبق معايرتها في إمكانية استخدام قيم المعلم للمفردات التي تم اشتغالها في الاختبار لحساب المنحنى المميز للاختبار ودالة معلومات الاختبار قبل تطبيق الاختبار . ويعتبر هذا أمراً ممكناً لأن أي منحنى من هذه المنحنيات يعتمد على توزيع درجات قدرة المفحوص في مقياس القدرة . ولذلك يمكن الحصول على كلا المنحنيين عندما تكون قيم معالم المفردات متاحة . وبافتراض هذين المنحنيين نجد أنه يكون لدى واضح الاختبار فكرة جيدة جداً عن كيفية أداء الاختبار قبل تقديمه لمجموعة من المفحوصين . وبالإضافة إلى ذلك وعند تطبيق ومعايرة الاختبار يمكن استخدام إجراءات معادلة الاختبار للتعبير عن تقديرات القدرة لمجموعة جديدة من المفحوصين في قياس قائمة المفردات .

بعض أهداف الاختبارات المقالية

Some Typical Testing Goals

وجعل تدريبات الحاسب مفهومة لك ، سيتم تعريف أنماط عديدة من أهداف الاختبار فيما يلي وستستخدم كأساس لأنماط اختبارات يتم تطويرها فيما بعد :

١- الاختبارات تصفية

لقد كان للاختبارات المستخدمة لأغراض التصفية القدرة على التمييز بصورة قاطعة بين المفحوصين ذوي القدرات أقل من مستوى قدرة معين والذين كانوا في ذلك

المستوى أو أعلى منه . ويتم استخدام هذه الاختبارات لتحسين المنح الدراسية وتحديد الطلاب لبرامج تعليمية خاصة مثل العلاج أو الوضع المتقدم .

٢- اختبارات المدى الواسع

ويتم استخدام هذه الاختبارات لقياس القدرة في مدى واسع من مقياس القدرة الأساسية . وقد كان الغرض الأول أن تكون قادراً على وضع مفردة عن قدرة المفحوص وتقران بين المفحوصين . وتعتبر اختبارات قياس القراءة أو الحساب أمثلة على اختبارات المدى الواسع .

٣- اختبارات الذروة (التمكن)

ويتم تصميم هذه الاختبارات لقياس القدرة بصورة جيدة تماماً في منطقة مقياس القدرة التي من الممكن أن توجد فيها معظم قدرات المفحوصين وبدرجة أقل خارج هذه المنطقة . وعندما يطور الشخص اختبار الذروة يمكنه حينئذ قياس القدرة بدرجة جيدة في مدى القدرة الأوسع من مدى اختبار التصفية لكنه ليس أوسع من اختبار المدى الواسع .

جلسة الحاسب الآلي للفصل الثامن

Computer Session for Chapter 8

الغرض من هذه الجلسة مساعدتك على تنمية القدرة لانتقاء المفردات من قائمة المفردات السابق معايرتها لتحقيق هدف الاختبار . وستقوم بوضع قيم معلم المفردات اختبار صمير لتحقيق أحد الأهداف الثلاثة للاختبار التي تم عرضها سابقاً . وسيتم بعد ذلك عرض المنحنى المميز للاختبار ودالة معلومات الاختبار على الشاشة ويمكنك تحديد فيما إذا ما تم تحقيق هدف الاختبار . وإن لم تكن كذلك يمكن انتقاء مجموعة جديدة من معالم المفردات والحصول على المنحنيات الناتجة . ويقليل من التدريب تصبح ذا كفاءة في وضع اختبارات ذات مزايا فنية متناسقة مع أهداف التصميم .

بعض القواعد الأساسية

Some Ground Rules

- ١- يفترض انتقاء المفردات على أساس المحتوى وقيم المعلم أيضاً وللأغراض الحالية ، ليس من الضروري إيضاح المحتوى الفعلي للمفردات .
- ٢- ليس لاي مفردتين في قائمة المفردات بالضبط نفس مجموعة قيم معلم المفردات .

٣- تعتبر قيم معلم المفردة للتقديرات التالية :

$$-3 \leq b \leq +3$$

$$0.50 \leq b < +2.00$$

$$0 \leq o \leq 0.35$$

وقد تم تقييد قيم معلم الميزة لكي تعكس مدى القيم التي نراها عادة في قائمة المفردات المحفوظة جيداً .

إجراءات لمثال حالة

Procedures for an Example Case

تقوم الآن بتشكيل اختبار تصفية ذي عشر مفردات حيث ستقوم بتقسيم المفحوصين إلى مجموعتين : الأولى تحتاج إلى تعليم علاجي والأخرى لا تحتاج إليه في القدرة التي تم قياسها بواسطة المفردات في قائمة المفردات . وسيحصل الطلاب الذين تقع قدرتهم أقل من قيمة

(١-) على توجيه وتعليم :

١- اتبع إجراءات البدء التي تم وصفها في المقدمة .

٢- استخدم الفأرة لتحديد TEST SPECIFICATION واضغط على

[SELECT] .

٣- اقرأ الشاشة التوضيحية ثم اضغط [CONTINUE] .

٤- اضغط على [NUMBER OF ITEMS] . وضع عدد المفردات في الاختبار إلى $N=10$.

٥- في قائمة SELECT ITEM CHARACTERISTIC CURVE MODEL ، اضغط على TWO PARAMETER .

٦- استجب لـ SETTINGS OK ؟ بالضغط على زر YES . وسيظهر على الشاشة ITEM PARAMETER .

٧- اضغط على [ENTER PARAMETERS] وضع بعد ذلك قيم بارامترات المفردات التالية :

المفردة	التمييز	الصعوبة
١	$a=1,2$	$b=-1,8$
٢	$a=1,4$	$b=-1,6$
٣	$a=1,1$	$b=-1,4$
٤	$a=1,3$	$b=-1,2$
٥	$a=1,5$	$b=-1,0$
٦	$a=1,0$	$b=-0,8$
٧	$a=1,4$	$b=-0,6$
٨	$a=1,2$	$b=-0,4$
٩	$a=1,1$	$b=-0,2$
١٠	$a=1,3$	$b=-0,0$

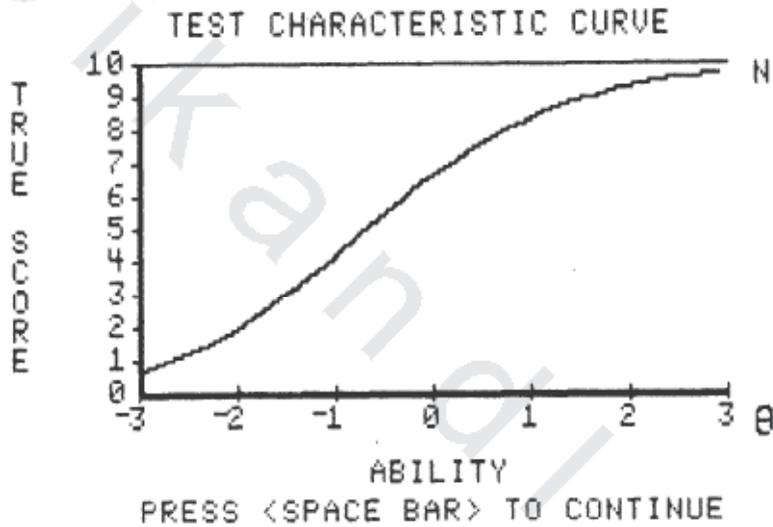
وقد كان المنطق الأساسي لهذه الاختيارات يشير إلى تمركز الصعوبات على مستوى نقطة القطع (١-) واستخدام القيم المعتدلة للتمييز .

٨- ادرس جدول معالم المفردة في لحظة . وإذا رغبت في تغيير القيمة ، اضغط على القيمة وسيظهر صندوق مدخل البيانات ويسمح لك بإدخال قيمة جديدة .

٩- وعندما تقبل بقييم المعلم أجب على رسالة PARAMETER VALUES OK ؟ بالضغط على زر YES .

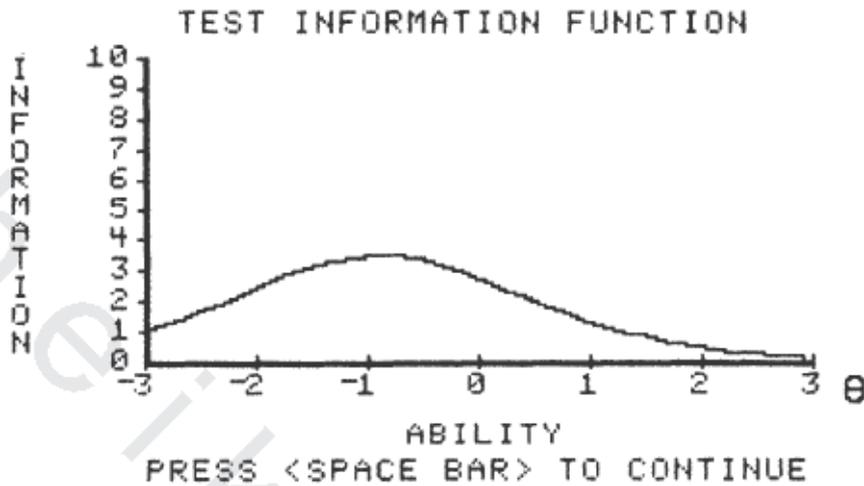
١٠- اضغط على [CONTINUE] سيظهر على الشاشة المنحنى المميز للاختبار
الموضح فيما يلي .

١١- وعندما يظهر على الشاشة المنحنى المميز للاختبار ، دون مستوى
القدرة التي تحدث عند درجة منتصف الدرجة الحقيقية . ودون أيضاً الحدار المنحنى في
ذلك المستوى للقدرة . وفيما يلي الشكل البياني :



الشكل رقم (١٠، ٨). المنحنى المميز للاختبار للمثال .

١٢- اضغط على [CONTINUE] وعندما يظهر على الشاشة دالة معلومات
الاختبار دون أقصى قدر من المعلومات ومستوى القدرة التي تحدث عندها . ويتم فيما
يلي إيضاح الدالة .



الشكل رقم (٨,٢). دالة معلومات الاختبار للمثال.

١٣- وإذا كنت تضغط على [DISPLAY FIRST CURVE] سيظهر المنحنى المميز للاختبار مرة أخرى. ولذلك يمكنك التبادل بين التين من الأشكال البيانية لدراسة علاقتها.

١٤- وقد كان هدف التصميم تحديد مفردات اختبار التصفية الذي سيؤثر في مستوى قدرة (-١,٠) وبصفة عامة نجد أنه تم تحقيق هذا الهدف. وقد تطابقت منتصف الدرجة الحقيقية The mid-true score مع مستوى قدرة (-١,٠) ولم يكن المنحنى المميز للاختبار مرتفعاً بصفة خاصة في نقطة القطع ويشير إلى أن الاختبار ينقصه التمييز. وقد حدثت فزوة معلومات الاختبار في مستوى قدرة (-١,٠) لكن الحد الأقصى كان صغيراً إلى حد ما. وتقترح النتائج أنه تم وضع الاختبار بدقة في مقياس القدرة لكي يمكن اكتشاف مجموعة أفضل من المفردات.

وستساعد التغيرات التالية على تحسين خصائص الاختبار: أولاً تجميع قيم صعوبات المفردة قريباً من نقطة القطع، ثانياً استخدام القيم الأكبر للمعالم التمييزية. وينبغي أن يساعد هذان التغيران على كون المنحنى المميز

للاختبار شديد الانحدار ويزيد من الحد الأقصى لقدر المعلومات في مستوى قدره (١,٠-).

١٥ - عندما تكون على شاشة دالة معلومات الاختبار اضغط [CONTINUE] وستظهر الشاشة التالية .

١٦ - أجب على سؤال DO ANOTHER TEST? بالضغط على زر YES .

١٧ - أجب على سؤال DO ANEW TEST? بالضغط على زر YES .

١٨ - أجب على سؤال PLOT ON SAME GRAPHS? بالضغط على زر

. YES

١٩ - أجب على سؤال SETTINGS OK ? بالضغط على زر YES .

٢٠ - أجب على سؤال PLOT ON SAME GRAPHS? بالضغط على زر

. YES

٢١ - كرر الخطوات من d إلى g باستخدام N=10 ومعالم المقررات التالية :

المقررة	التمييز	الصعوبة
١	$\sigma=1,9$	$b=1,1$
٢	$\sigma=1,7$	$b=1,0$
٣	$\sigma=1,8$	$b=1,1$
٤	$\sigma=1,6$	$b=1,2$
٥	$\sigma=1,9$	$b=1,0$
٦	$\sigma=1,8$	$b=0,8$
٧	$\sigma=1,9$	$b=0,9$
٨	$\sigma=1,9$	$b=1,0$
٩	$\sigma=1,7$	$b=0,9$
١٠	$\sigma=1,6$	$b=1,0$

- ٢٢- وعندما يظهر المنحنى المميز للاختبار ، قم بمقارنته بالمنحنى السابق الذي لا يزال على الشاشة ، ثم حدد إذا تم زيادة المنحدر المنحنى في مستوى قدره (-١,٠) .
- ٢٣- وعند ما تظهر دالة معلومات الاختبار قم بمقارنتها بالدالة الموجودة على الشاشة ثم حدد إذا كان أقصى قدر من المعلومات أكبر مما كان في مستوى قدرة (-١,٠) .
- ٢٤- وإذا سارت الأمور كلها على ما يرام ، ضع مفردات اختبار جديد لتحسين الخصائص الفنية للاختبار كما تم إيضاحه في المنحنى المميز للاختبار ودالة معلومات الاختبار .

التعريفات

وفي كل التعريفات التالية ضع عدداً من معالم المفردة . وعندما ترى المنحنى المميز للاختبار ودالة معلومات الاختبار ، استخدم ملمح التحرير لتغيير قيم معلم المفردات . ودع المنحنيات الجديدة مع المنحنيات السابقة . وستسمح لك هذه الإجراءات لرؤية تأثير التغييرات . ثم كرر هذه العملية حتى تشعر أنك قد أجزت هدف تخصيص الاختبار .

التعريف الأول

قم بتكوين اختبار التصفية ذي العشر مفردات للأداء في مستوى قدرة (+٠,٧٥) باستخدام نموذج راش .

التعريف الثاني

قم بتكوين اختبار المدى الواسع في ضوء النموذج الثلاثي المعلم الذي سيكون له دالة معلومات الاختبار أفضياً في مدى قدرة من (-١,٠) إلى (+١,٠)

التعريف الثالث

قم بتكوين اختبار ذي منحنى يميز للاختبار بالمحدد صغير ودالة معلومات ذات شكل دائري بصورة معتدلة . استخدم إما النموذج الثاني وإما الثلاثي المعلم .

تدريب الرابع

قم بتكوين الاختبار الذي سيكون له منحنى يميز للاختبار الخطي تقريباً والذي تحدث فيه منتصف الدرجة الحقيقية عند مستوى القدرة الصفر . استخدم نموذج راس .

تدريب الخامس

كرر المسألة السابقة باستخدام النموذج الثلاثي المعلم .

تدريب السادس

قم بتكوين اختبار ذي دالة معلومات أفقية ومدى القدرة من (-٢,٠) إلى (+٢,٠) بأقصى قدر من المعلومات يصل إلى (٢,٥) .

تدريب السابع

استخدم جلسة الحاسب لتجربة المجموعات المختلفة لأهداف الاختبار ، نماذج المنحنى المميز للمفردات ، وعدد المفردات . والهدف هو إمكانية الحصول على المنحنيات المميزة للاختبار ، ودالة معلومات الاختبار المثلى لأهداف الاختبار . وسيكون من المفيد استخدام ملمح التحرير لتغيير قيم معلم المفردة الخاصة بدلاً من إعادة إدخال مجموعة كاملة من قيم معلم المفردات لكل محاولة .

ملاحظات Things To Notice

١- الخصائص الصفيية

أ) كان للمنحنى المميز للاختبار المرغوب فيه منتصف الدرجة الحقيقية عند نقطة القطع مستوى قدرة نوعي ، وينبغي أن يكون المنحنى شديد الانحدار بقدر الإمكان كما كان في مستوى القدرة .

ب) ينبغي أن تبلغ دالة معلومات الاختبار اللزوة بأقصى مداها عند نقطة قطع مستوى القدرة.

ج) . ينبغي أن تكون قيم معالم صعوبة المفردة متجمعة قدر الإمكان حول نقطة القطع المرغوبة للقدرة . وتعتبر الحالة المثلى وجود كل قيم صعوبة المفردة عند درجة القطع وتميزات المفردة كبيرة . ومع ذلك يعتبر ذلك غير واقعي لأن قائمة المفردات نادراً ما تحتوي على مفردات كافية ذات قيم صعوبة مشتركة . لو ان الاختيار من بين المفردات فقم باختيار المفردات التي تنتج أقصى قدر من المعلومات عند نقطة القطع .

٢- الخصائص الممدى الواسع

أ) لقد كان للمنحنى المميز للاختبار المرغوب له منتصف درجة حقيقية عند مستوى قدرة تناظر نقطة متوسط مدى القدرة المرغوبة . وكان ذلك في أغلب الأحيان مستوى قدرة الصفر . وينبغي أن يكون المنحنى المميز للاختبار خطأً مستقيماً معظم المدى .

ب) تعتبر دالة معلومات الاختبار المرغوب فيها أفضى في أوسع مدى ممكن . وينبغي أن يكون أقصى قدر من المعلومات كبيراً بقدر الإمكان .

ج) وينبغي أن تنتشر قيم معالم صعوبة المفردة باتساق على مقياس القدرة وأن تكون عملية قدر الإمكان . ويوجد تعارض بين أهداف أقصى قدر من المعلومات ودالة معلومات الاختبار الأفضى . ولتحقيق دالة معلومات الاختبار الأفضى يعتبر من الضروري أن يكون للمفردات ذات التمييز من المنخفض للمعتدل التي لها توزيع شكل U بصعوبات المفردات . ومع ذلك تنتج هذه المفردات قدر أقل من المعلومات وستكون الدقة الكلية منخفضة .

٣- الخصائص الفردية

أ) وقد كان للمنحنى المميز للاختبار المرغوب فيه درجة متوسطة صحيحة عند مستوى قدرة في وسط مدى القدرة المرغوبة . وينبغي أن يكون للمنحنى المنحدر في مستوى القدرة .

ب) ينبغي أن يكون لدالة معلومات الاختبار المرغوب فيه أقصى حد لها وفي نفس مستوى القدرة كمنتصف للمنحنى درجة الحقيقية المميز للاختبار وينبغي أن تكون دالة معلومات الاختبار دائرية في شكلها في مدى القدرة المرغوبة .

ج) تجميع صعوبات المفردات حول متوسط مدى القدرة المرغوبة لكن لا تكون بإحكام كما في حالة اختبار التصفية . وينبغي أن تكون قيم المعالم التمييزية كبيرة بصفة عملية . وينبغي أن يكون للمفردات ذات الصعوبات في مدى القدرة المرغوبة قيم أكبر للتمييز عن المفردات ذات الصعوبات خارج هذا المدى .

٤- دور نماذج المنحنى المميز للمفردة

أ) وبسبب قيمة المعالم التمييزية التي كانت ثابتة في (١,٠) فقد كان لنموذج راش حد لأقصى قدر من المعلومات التي يمكن الحصول عليها ، وهو (٠,٢٥) حيث $P_1(\hat{\theta})=0.5$ عندما $Q_1(\hat{\theta})=0.25$ ولذلك كان أقصى قدر نظري من المعلومات (٠,٢٥) مرة من عدد المفردات .

ب) في حالة وجود معلم التخمين سيُنتج النموذج ثلاثي المعلم منحنى مميزاً للاختبار الخطي ودالة معلومات الاختبار ذات المستوى العام الأكثر انخفاضاً مما كان في النموذج ثنائي المعلم ذو نفس مجموعة معالم الصعوبة والتمييز . وقد كانت دالة

المعلومات في النموذج ثنائي المعلم الرابطة العليا لدالة المعلومات في النموذج ثلاثي المعلم عندما تكون قيم p, q نفس الشيء * .

(ج) ولأغراض تخصيص الاختبار يفضل الباحث النموذج ثنائي المعلم .
 • - دور عدد المقدرات

(أ) إن زيادة عدد المقدرات له تأثير قليل على الشكل العام للمنحنى المميز للاختبار إذ ظل توزيع مجموعة معالم المقدرات نفس الشيء * .

(ب) إن زيادة عدد المقدرات في الاختبار له تأثير دال على المستوى العام لدالة معلومات الاختبار . وكان الموقف الأمثل وجود عدد كبير من المقدرات ذات القيم المرتفعة للمعالم التمييزية وتوزيع صعوبات المقدرات المتناسقة مع أهداف الاختبار .

(ج) ويعتبر الأسلوب الذي تم فيه توليف قيم معالم المقدرات ذو اعتبار هام . وعلى سبيل المثال اختيار القيم المرتفعة لمؤشر التمييز للمقدرات التي لم تكن صعوبتها ذات أهمية فيما يتعلق بدالة معلومات الاختبار أو انحدار المنحنى المميز للاختبار . ولذلك يجب أن يتصور واضح الاختبار ما يشكله المنحنى المميز للاختبار ودالة معلومات الاختبار للتأكد من إسهام المقردة في المنحنى المميز للاختبار ودالة معلومات الاختبار .

المراجع

References

- Birnbaum, A. "Some latent trait models and their use in inferring an examinee's ability." Part 5 in F.M. Lord and M.R. Novick. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley, 1968.
- Hambleton, R.K., and Swaminathan, H. *Item Response Theory: Principles and Applications*. Hingham, MA: Kluwer, Nijhoff, 1984.
- Hulin, C. L., Drasgow, F., and Parsons, C.K. *Item Response Theory: Application to Psychological Measurement*. Homewood, IL: Dow-Jones, Irwin: 1983.
- Lord, F.M. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Erlbaum, 1980.
- Mislevy, R.J., and Bock, R.D. *PC-BILOG 3: Item Analysis and Test Scoring with Binary Logistic Models*. Mooresville, IN: Scientific Software, Inc, 1986.
- Wright, B.D., and Mead, R.J. BICAL: Calibrating Items with the Rasch Model. Research Memorandum No. 23. Statistical Laboratory, Department of Education, University of Chicago, 1976.
- Wright, B.D., and Stone, M.A. *Best Test Design*. Chicago: MESA Press, 1979.
- NOTE: See (www.assess.com) for more information about the BILOG program. See (www.winsteps.com) for more information about BICAL and its successors, WINSTEPS and BIGSTEPS.

Selected Resources on Item Response Theory Web Sites and Online Resources

Cumulative Item Response Theory Models

<http://www.education.umd.edu/Depts/EDMS/tutorials/CIRT.html>

Currently under development by the University of Maryland's Department of Measurement, Statistics, and Evaluation, this Web site will provide IRT models for binary and polytomous responses. Parametric and nonparametric models and applications will be addressed. ERIC Clearinghouse on Assessment and Evaluation

<http://ericae.net>

The ERIC Clearinghouse on Assessment and Evaluation, hosted by the University of Maryland and funded by the U.S. Department of Education, maintains an extensive body of resources on the Web, including an online refereed journal called *Practical Assessment, Research & Evaluation* and a full-text library with more than 400 key reports from respected organizations.

Institute for Objective Measurement

www.rasch.org

The host of this site draws a careful distinction between IRT and Rasch measurement, but those interested in IRT will likely find items of interest here, including the full text of *Rasch Measurement Transactions*, the quarterly publication of the Rasch Measurement SIG of the American Educational Research Association.

Interactive CAT & IRT Mini-Tutorial

<http://ericae.net/scripts/cat/catdemo.htm>

Part of an online, interactive tutorial on computer adaptive testing developed, this mini-tutorial introduces the three-parameter IRT model and allows users to experiment with varying the item parameter values and generating graphs of item response functions.

Item Response Theory Models for Unfolding

<http://www.education.umd.edu/EDMS/tutorials/Intro.html>

This Web site, maintained by James S. Roberts, Assistant Professor in the Department of Measurement, Statistics and Evaluation at the University of Maryland, explores unfolding models for predicting item scores where responses are obtained in a rating scale format such as a Thurstone or Likert scale. Users can download free modeling software and illustrative data sets on attitudes toward capital punishment or censorship.

Electronic Discussion Groups

Rasch Discussion Listserv

This unmoderated forum sponsored by the Australian Council for Educational Research has operated since 1966 to support "the exchange of news, questions and answers about the theory and practice of Rasch Measurement."

To join, send an e-mail with text *subscribe rasch* to mailserv@acer.edu.au.

ERIC Database Search

Abstracts for several hundred documents and journal articles pertaining to item response theory have been indexed in the ERIC database. You can conduct your own search of ERIC for these citations via the Internet by using our online thesaurus-driven search engine, which we call the ERIC Search Wizard. Go to:

<http://ericse.net/>

Choose "Search ERIC"

You'll be taken to <http://searcheric.org>

Enter the term "item response theory" into the search box.

On the right side of the screen, you'll see a definition of the term and a list of related terms. Start your search by clicking on the box in front of Item Response Theory and then scroll down to the bottom of the column, where you will click on "Add to Set 1."

Use the related terms or the Look Up box to enter additional terms to combine with Item Response Theory in order to narrow your search.

You may, for example, try Test Construction, Test Items, Item Bias, Item Analysis, Adaptive Testing, Test Reliability, or Test Validity in Set 2 or Set 3.

Print Classics

Applications of Item Response Theory to Practical Testing Problems

F. M. Lord

Lawrence Erlbaum, 1980

This classic text offers a thorough technical presentation of IRT models, including their limitations, as well as discussion of such practical problems as estimating ability and item parameters, equating, study of item bias, omitted responses and formula scoring. Flexilevel tests, multilevel tests, tailored testing, and mastery testing are also addressed.

Fundamentals of Item Response Theory

R. K. Hambleton, H. Swaminathan, and H. J. Rogers

Sage, 1991

This introductory text draws on concepts from classical measurement methods and basic statistics to present the basics of IRT and its applications in test construction, identification of potentially biased test items, test equating, and computerized adaptive testing. Alternative procedures for estimating IRT parameters, including maximum likelihood estimation, marginal maximum likelihood estimation, and Bayesian estimation, are discussed. Step-by-step numerical examples are included throughout.

Item Response Theory: Parameter Estimation Techniques

F.B. Baker

Marcell Dekker, 1992

This book presents the mathematical details of the parameter estimation procedures used in item response theory. The procedures maximum likelihood, marginal maximum likelihood, and Bayesian estimation are presented for binary, graded, and nominal response items. BASIC computer programs for these procedures are provided in the book.

Handbook of Modern Item Response Theory

W. J. van der Linden and R. K. Hambleton, eds.

Springer, 1997

This reference work provides an introduction to item response theory and its application to educational and psychological testing. A comprehensive treatment of models and families of models is provided in 27 chapters, each of which is authored by person(s) who either proposed or contributed substantially to the development of the model discussed. Each chapter includes an introduction, presentation of the model, parameter estimation and goodness of fit, and a brief empirical example. Some chapters also offer discussion.

ثبت المصطلحات

أولاً: عربي - إنجليزي

1

incorrect response	الإجابات الخاطئة
curve-fitting procedure	إجراءات مطابقة المنحنى
probability of correct response	احتمالية الاستجابة الصحيحة
screening test	اختبار التصفية
peaked test.	اختبار الذروة
broad-range tests	اختبارات المدى الواسع
anchor the metric	الإرساء المترى
correct response	الاستجابات الصحيحة
alternate forms	الأشكال المتتالية
maximum value of a true score	أعلى قيمة للدرجة الحقيقية
maximum likelihood	أقصى ترجيح
maximum amount of information	أقصى مقدار للمعلومة

logistic deviate

الانحراف اللوجستي

١

item banking

بنك الأسئلة

BICAL

بيكال

٢

linear transformation

التحويل الخطي

dichotomously scored

التسجيل الثنائي للدرجات

interpreting the test information function

تفسير دالة معلومة الاختبار

maximum likelihood estimation

تقدير الترجيح الأقصى

estimating item parameters.

تقدير معالم المفردة

discrimination

التمييز

perfect discrimination

التمييز التام

negative discrimination

التمييز السلبي

٣

group invariance of item parameters

ثبات المجموعة لمعالم المفردة

٤

compute the points on an item characteristic curve

حساب النقاط على منحنى خصائص المفردة

المفردة

٥

estimated standard error

الخطأ المعياري المقدر

standard error of estimate

الخطأ المعياري للتقدير

logistic function	الدالة اللوجستية
information function	دالة المعلومة
ideal information function	دالة المعلومة المثالية
$P(\cdot)$	الدالة سينا
ideal test information function	دالة معلومة الاختبار المثالية
item information function	دالة معلومة المفردة
test information function	دالة معلومة الاختبار
raw test score	درجة الاختبار الخام
perfect test score	درجة الاختبار الكاملة
raw score	الدرجة الخام
ability score	درجة القدرة
precision	الدقة
location index	دليل الموضع
increasing the number of items	زيادة عدد المفردات
latent trait	السمة الكامنة
item difficulty of classical theory	صعوبة المفردة وفقاً للنظرية التقليدية
row marginals	صف الهوامش

anchoring process

عملية الإرساء

precalibrated item pool

قائمة المفردة قبل المعايرة

ability

القدرة

item pools

قوائم المفردة

metric

القياسي

initial values

القيم الأولية

item parameter values

قيم معلم المفردة

invariance principle

مبدأ الثبات

normal ogive

متجمع طبيعي

computational demands

متطلبات حسابية

symmetric

المتماثل

edited data set

مجموعة البيانات المحررة

ability score group

مجموعة درجة القدرة

ability level

مستوى القدرة

test equating

معادلة الاختبار

test calibration

معايرة الاختبار

guessing parameter

معلم التخمين

discrimination parameter

معلم التمييز

difficulty parameter	معلم الصعوبة
ability parameter	معلم القدرة
test information	معلومة الاختبار
item information	معلومة المفردة
binary items	المفردات المزدوجة
amount of information	مقدار المعلومات
ability scale	مقياس القدرة
scale of measurement	مقياس تدرج القياس
mid-trus score	متصف الدرجة الحقيقية
test characteristic curve	منحنى خصائص الاختبار
item characteristic curve	منحنى خصائص المفردة
test constructor	موجة الاختبار
item response vectors	موجهات الاستجابة للمفردة
observed proportion of correct response	النسبة الملاحظة للإجابة الصحيحة
classical test theory	نظرية الاختبارات التقليدية
item response theory	نظرية الاستجابة للمفردة
itemized theory	نظرية تفضيل المفردات
mathematical models	النماذج الرياضية
one-parameter model	النموذج أحادي المعلم
logistic model	النموذج اللوجستي

one-parameter logistic model	النموذج اللوجستي أحادي المعلم
normal ogive model	النموذج المتجمع الطبيعي
three-parameter model	النموذج ثلاثي المعلم
two-parameter model	النموذج ثنائي المعلم
rasch model	نموذج راش
item characteristic curve model	نموذج منحنى خصائص المفردة
column marginals	هوامش الأعمدة
LOGIST	وحدة اللوجست
logit	وحدة اللوجيت

ثانياً: إنجليزي - عربي

A

ability	القدرة
ability level	مستوى القدرة
ability parameter	معلم القدرة
ability scale	مقياس القدرة
ability score	درجة القدرة
ability score group	مجموعة درجة القدرة
alternate forms	الأشكال المتتالية
amount of information	مقدار المعلومات
anchor the metric	الإرساء المترى
anchoring process	عملية الإرساء

B

BICAL	بيكال
binary items	المفردات المزدوجة
broad-range tests	اختبارات المدى الواسع

C

classical test theory	نظرية الاختبارات التقليدية
column marginals	هوامش الأعمدة
computational demands	متطلبات حسابية
compute the points on an item characteristic curve	حساب النقاط على منحنى خصائص المفردة

correct response

الاستجابات الصحيحة

curve-fitting procedure

إجراءات مطابقة المنحنى

D

dichotomously scored

التسجيل الثنائي للدرجات

difficulty parameter

معلم الصعوبة

discrimination

التمييز

discrimination parameter

معلم التمييز

E

edited data set

مجموعة البيانات المحررة

estimated standard error

اخطأ المعياري المقدر

estimating item parameters.

تقدير معالم المفردة

group invariance of item parameters

ثبات المجموعة لمعالم المفردة

guessing parameter

معلم التخمين

I

ideal information function

دالة المعلومة المثالية

ideal test information function

دالة معلومة الاختبار المثالية

incorrect response

الإجابات الخاطئة

increasing the number of items

زيادة عدد المفردات

information function

دالة المعلومة

initial values

القيم الأولية

interpreting the test information function	تفسير دالة معلومة الاختبار
invariance principle	مبدأ الثبات
item banking	بنك الأسئلة
item characteristic curve	منحنى خصائص المفردة
item characteristic curve model	نموذج منحنى خصائص المفردة
item difficulty of classical theory	صعوبة المفردة وفقاً للنظرية التقليدية
item information	معلومة المفردة
item information function	دالة معلومة المفردة
item parameter values	قيم معلم المفردة
item pools	قوائم المفردة
item response theory	نظرية الاستجابة للمفردة
item response vectors	موجهات الاستجابة للمفردة
itemized theory	نظرية تفضيل المفردات
latent trait	السمة الكامنة
linear transformation	التحويل الخطي
location index	دليل الموضع
LOGIT	وحدة اللوجست
logistic deviate	الانحراف اللوجستي
logistic function	الدالة اللوجستية
logistic model	النموذج اللوجستي

logit

وحدة اللوجيت

M

mathematical models

النماذج الرياضية

maximum amount of information

أقصى مقدار للمعلومة

maximum likelihood

أقصى ترجيح.

maximum likelihood estimation

تقدير الترجيح الأقصى

maximum value of a true score

أعلى قيمة للدرجة الحقيقية

metric

القياسي

mid-true score

متصف الدرجة الحقيقية

N

negative discrimination

التمييز السلبي

normal ogive

متجمع طبيعي

normal ogive model

النموذج المتجمع الطبيعي

O

observed proportion of correct response

النسبة الملاحظة للإجابة الصحيحة

one-parameter logistic model

النموذج اللوجستي أحادي المعلم

one-parameter model

النموذج أحادي المعلم

P

P (θ)

الدالة سيتا

peaked test.

اختبار الذروة

perfect discrimination

التمييز التام

perfect test score	درجة الاختبار الكاملة
precalibrated item pool	قائمة المفردة قبل المعايرة
precision	الدقة
probability of correct response	احتمالية الاستجابة الصحيحة
R	
rasch model	نموذج راش
raw score	الدرجة الخام
raw test score	درجة الاختبار الخام
row marginals	صف الهوامش
S	
scale of measurement	مقياس تدرج القياس
screening test	اختبار التصفية
standard error of estimate	خطأ المعياري للتقدير
symmetric	المتماثل
T	
test calibration	معايرة الاختبار
test characteristic curve	منحنى خصائص الاختبار
test constructor	موجة الاختبار
test equating	معادلة الاختبار
test information	معلومة الاختبار
test information function	دالة معلومة الاختبار

three-parameter model

النموذج ثلاثي المعلم

two-parameter model

النموذج ثنائي المعلم

obeikandi.com

كشاف الموضوعات

أقصى مقدار للمعلومة ٩٨	١
الانحراف اللوجستي ٢٦	الإجابات الخاطئة ٢٨
بنك الأسئلة ١٤٧	إجراءات مطابقة المنحنى ٥٠
بيكال ١٢٦ ، ١٢٩	احتمالية الاستجابة الصحيحة ٢٤ ، ٧٩
التحويل الخطي ١٢٥	اختبار التصفية ١٥٠ ، ١٥٤ ، ١٥٨ ، ١٥٦
التسجيل الثنائي للدرجات ١٢٦	اختبار النزوة ١٤٩ ، ١٥٠
تفسير دالة معلومة الاختبار ١٣٣ ، ١٣٥ ، ١٣٩ ، ١٤٤	اختبارات المدى الواسع ١٤٩
تقدير الترجيح الأقصى ٤٢ ، ٤٩	الإرساء المتري ١٢٨
تقدير معالم المفردة ٤١ ، ٤٤	الاستجابات الصحيحة ٣ ، ٤ ، ٥ ، ٢٨ ، ٢٣ ، ٦
التمييز ٤ ، ٦ ، ٧	الأشكال المتتالية ١٥٤
التمييز التام ٧	أعلى قيمة للترجة الحقيقية ١٣٠
التمييز السلبي ٢٧ ، ٧١	أقصى ترجيح ١٢٤

- درجة القدرة ٢، ٧٧
- الدقة ١٠٨، ١١٨، ١٥٨
- دليل الموضوع ٥
- د
- زيادة عدد المقدرات ١٥٩
- س
- السمة الكامنة ٤٩، ٩٩، ١٣٢،
- ١٤٨، ١٣٣
- ص
- صعوبة المفردة وفقاً للنظرية
- التقليدية ٧٥
- صف الهوامش ٢٨
- ع
- عملية الإرساء ١٢٨
- ح
- قائمة المفردة قبل المعايرة ١٤٨
- القدرة ٢٥، ٢٦، ٢٧، ٢٨
- قوائم المفردة ١٤٧
- القياسي ٢
- القيم الأولية ٧٨
- قيم معلم المفردة ١٣٠، ١٤٩
- ثبات المجموعة لعالم المفردة ٤٨
- ح
- حساب النقاط على منحني
- خصائص المفردة ١٩
- أ
- الخطأ المعياري المقدر ٨١
- الخطأ المعياري للتقدير ٩١، ٩٣
- د
- الدالة اللوجستية ١٨
- دالة المعلومة ٩٨
- دالة المعلومة المثالية ٩٨
- الدالة سيتا ٣
- دالة معلومة الاختبار المثالية ١٥٣،
- ١٥٧
- دالة معلومة المفردة ٩٩
- دالة معلومة الاختبار ١٥٦، ١٥٩
- درجة الاختبار الختام ٣، ٥٩
- درجة الاختبار الكاملة ١٠٠
- الدرجة الختام ١٢٧، ١٣٠،
- ١٣١، ١٣٦، ١٣٨

مقياس القدرة ٦، ١٩، ١٢٣،

١٤٨

مقياس تدرج القياس ٣٨

متصف الدرجة الحقيقية ٣٠

منحنى خصائص الاختبار ١٠٧

منحنى خصائص المفردة ١٠٢،

١٠٤، ١٠٣

موجة الاختبار ١٣٠

موجهات الاستجابة للمفردة ١٤٧

٥

النسبة الملاحظة للإجابة

الصحيحة ٥٢

نظرية الاختبارات التقليدية ١٤٧

نظرية الاستجابة للمفردة ٤٥، ٥٠

نظرية تفضيل المفردات ١٥

النماذج الرياضية ١٧

النموذج أحادي المعلم ٢١، ٢٣،

٢٤، ٣٩، ٤٤، ٥٥

النموذج اللوجستي ٢١، ٢٢

٥

مبدأ الثبات ٥٤

متجمع طبيعي ٣٩

متطلبات حسابية ٤٤

المتماثل ١٠٣، ١١٣

مجموعة البيانات المخررة ١٣٤

مجموعة درجة القدرة ٤٤

مستوى القدرة ٤، ٦، ١٩، ٢٥،

٦٠

معادلة الاختبار ١٤٠، ١٤٤

معايرة الاختبار ١٢٤، ١٢٦،

١٢٧، ١٣٢

معلم التخمين ٢٥

معلم التمييز ٢٥

معلم الصعوبة ٢٥

معلم القدرة ٩٧، ١٢٩

معلومة الاختبار ١٠٦

معلومة المفردة ١٠٦

المفردات المزدوجة ١٣

مقدار المعلومات ١٠٤، ١٠٥

د

هوامش الأعمدة ١٢٨

هـ

وحدة اللوجيست ١٨ ، ٢٣ ، ٢٦

وحدة اللوجيست ٢٣

النموذج اللوجستي أحادي

المعلم ٢١

النموذج المتجمع الطبيعي ٢٥

النموذج ثلاثي المعلم ٢٦

النموذج ثنائي المعلم ٣٩

نموذج راش ٢١ ، ٢٢ ، ٢٤ ، ٣٣ ،

٣٩

نموذج منحنى خصائص المفردة ٧٩