

الملحق أ

برامج الحاسب المستخدمة في الاختبارات والامتحانات^(١)

المقدمة

يعتمد تشغيل البرامج الموجودة على هذا القرص المرن على qbasic أو basica يمكن تحميل qbasic من القرص، ويمكن تشغيل البرامج من القرص المرن أو من القرص الصلب (وغالباً ما تكون البرامج أسرع عند تشغيلها على القرص الصلب). ينصح باتباع التعليمات التالية لنسخ الملفات من القرص المرن إلى القرص الصلب:

الحد الأدنى لمتطلبات نظام التشغيل

- IBM-PC أو معادلاته من الأجهزة التي تضم نسخة DOS 3.3 أو أعلى.
- محرك قرص مرن ٣,٥ إنش.

كيفية تثبيت الملفات على الحاسب

تصحك بنسخ الملفات على القرص الصلب وذلك بتشغيل برنامج التثبيت. ولتثبيت الملفات، إتبع الخطوات التالية:

- ١- أدخل القرص في محرك القرص المرن.
- ٢- إطبّع A:\INSTALL عند ظهور أمر DOS، اضغط مفتاح الإدخال. ستظهر شاشة تثبيت البرنامج، اضغط مفتاح الإدخال للتتابع.
- ٣- إن الفهرس الذي يظهر تلقائياً لتخزين الملفات هو C:\TESTS، ويمكن تغيير هذا الموقع إذا شئت. نفذ التعليمات التي تظهر على الشاشة.

(١) يعتمد تشغيل هذه البرامج على qbasic و basica. يتوفر على القرص المرن الملف التنفيذي qbasic.exe والذي وفرته لنا شركة مايكروسوفت. جميع الحقوق محفوظة لشركة مايكروسوفت.

٤- سيقوم برنامج التثبيت بنسخ كافة الملفات بما فيها برنامج qbasic إلى قرصك الصلب في فهرس C:\TEST أو أي فهرس آخر يحدده المستخدم.

كيفية استخدام البرامج

عند تشغيل البرامج من القرص المرن قم بعمل التالي:

١- أدخل القرص المرن في مشغل الأقراص المرنة في جهاز الحاسب.

٢- إطبغ A: في DOS عند ظهور أمر DOS

٣- اطبغ menu أو prog في سطر >A:\ ثم اضغط مفتاح الإدخال. الأمر menu سوف يدخلك إلى قائمة البرامج والتي يمكنك من اختيار البرنامج الذي تريد تشغيله. يتبع الأمر prog السؤال: "category and number of program?" والتي يتوجب عليك تحديد فئة لها من (A إلى H) ورقم البرنامج الذي ترغب بتشغيله مثلاً (prog.al).

وبإمكانك كذلك تشغيل البرنامج وأنت في Windows من قائمة File في program manager أو في قائمة start اختر الأمر Run اطبغ a: menu أو a: prog في حقل التعليمات ثم اضغط الزر Ok.

إذا قمت بتثبيت البرنامج على قرص الحاسب الصلب قم بعمل التالي:

١- في Dos اطبغ cd c:\tests (أو الفهرس الذي اختاره المستخدم) واضغط على مفتاح الإدخال.

٢- اطبغ menu أو prog في سطر >c:\ tests واضغط مفتاح الإدخال. الأمر menu سوف يدخلك على قائمة البرامج والتي بإمكانك اختيار البرنامج الذي تريد تشغيله منها. يتبع الأمر prog السؤال: "category and number of program?" مما يتطلب منك تحديد فئة البرنامج (من A إلى H)، بإمكانك الخروج من أي برنامج قيد التشغيل بضغط المفتاح F1.

إن مخرجات بعض البرامج سوف تظهر على الشاشة أمامك و/أو في ملف مخرجات على القرص يحمل اسم results. اتبع التعليمات لكل برنامج بحذر ولن تحصل معك مشاكل تذكر. وتذكر دائماً أن الامتحانات المدرجة في بعض البرامج يقصد بها أن تكون تمارين توضيحية أو للعرض فقط أكثر مما هي اختبار جدي بهدف التعليم.

لا تتوفر معايير تمثيلية لعينات الاختبارات الموجودة، ولذلك فمن الأفضل استعراض النتائج على أنها نتائج مقترحة وليست نهائية أو حاسمة. مساعدة المستخدم

إذا احتجت لمساعدة مبدئية في عملية التثبيت أو اضطررت لاستبدال قرص مرن نالف، يرجى الاتصال بقسم دعم المنتجات على الرقم ٦٧٥٣-٨٥٠ (٢١٢) خلال أيام الأسبوع من التاسعة صباحاً وحتى الرابعة بعد الزوال بتوقيت شرق أمريكا.

للقيام بطلبات إضافية أو لطلب معلومات إضافية عن منتجات Wiley، يرجى الإتصال بـ ٥٩٤٥ ٢٢٥ (٨٠٠).

معلومات عن البرنامج

فيما يلي وصف مختصر قد يساعدك في التعرف على البرامج:

برامج الفئة A: الطرق الأساسية في الإحصاء

١- التوزيعات التكرارية و الرسوم البيانية المرتبطة بها: يقوم هذا البرنامج بإنشاء التوزيعات التكرارية للدرجات وإعداد المدرجات التكرارية والمضلعات لمجموعة من الدرجات. يقوم المستخدم بتحديد عدد المجالات التي يرغب بها وفيما إذا كان يحتاج لمدرج أولمضلع التكرار. وتظهر على الشاشة، بالإضافة إلى المجالات والتكرارات المقابلة لها، حدود الدرجات العليا والدنيا.

٢- الإحصاء الوصفي: بناءً على مجموعة من الدرجات غير المصنفة والتي لا تتجاوز ١٠٠ درجة، يقوم هذا البرنامج بحساب وطبع الحدود العليا والدنيا للدرجات، ومدى الدرجات، والوسط الحسابي، والوسيط، والتباين، والانحراف المعياري، ومؤشرات الإلتواء والتفرطح، والدرجات الخام والدرجات Z الموافقة لها:

٣- برنامج رسم مخطط التبعثر وخط الانحدار: يقوم هذا البرنامج برسم مخطط التبعثر وخط الانحدار لـ X على Y ولـ X على Y لجملة من معدلات الدرجات. كما يقوم بحساب وطباعة القيم العددية للمنحدر وبتوقع مجال التقاطع لتنبؤ X من Y و Y من X ومعامل الارتباط المولد للعزوم (r) .

٤- الانحدار والارتباط الخطي البسيط: يقوم هذا البرنامج بحساب معامل الارتباط المولد للعزوم (r) بين X و Y ، كما يقوم بحساب معادلة الانحدار الخطي للتنبؤ بـ Y من X والأوساط الحسابية والانحرافات المعيارية لـ X و Y والخطأ المعياري للتقدير وقيمة Z لاختبار للدلالة الإحصائية لمعامل الارتباط لـ $3 - 100$ زوج من قيم X و Y . كما يمكن تحديد عدد قيم الأخطاء السلبية والإيجابية غير الحقيقية وعدد حالات القبول والرفض للممتحنين الذين وقعت درجاتهم دون الحد الفاصل للدرجات (أو دون مستوى الأداء المقبول).

٥- تحليل الانحدار المتعدد: يقوم هذا البرنامج بحساب أثقال الانحدار القياسي وغير القياسي ومعامل الارتباط المتعدد (R) والأخطاء المعيارية لأوزان الانحدار وقيم التحديد للدلالة الإحصائية لأوزان الانحدار لتحليل إنحدار خطي بالنسبة لمتغير مستقل واحد أو لمتغيرين مستقلين أو لثلاثة متغيرات مستقلة. والمتغير الأخير هو متغير تابع. المدخلات هي الأوساط الحسابية والانحرافات المعيارية والارتباطات الداخلية بين المتغيرات.

برامج الفئة B: وضع وإجراء وتصحيح الاختبارات

١- بناء الاختبارات الموضوعية: يسمح هذا البرنامج للمستخدم بالقيام بما يلي:

• بناء اختبارات موضوعية للشخصية

• استعراض ومراجعة الاختبار الذي تم بناؤه

• وضع مفتاح الإجابات للاختبار

• استعراض ومراجعة مفتاح إجابات الاختبار

• طباعة الاختبار

• طباعة مفتاح إجابات الاختبار

• تغيير كلمة سر البرنامج

• إدخال سلسلة من أرقام التعريف (الأرقام الامتحانية التي يحملها كل من يتقدم للاختبار) للاحتفاظ بسجل لمن يتم اختباره ولضمان أن يتم اختبار كل فرد مرة واحدة فقط.

إن كلمة السر الأولية للبرنامج هي Makeit. يتم تخزين أرقام التعريف من خلال الاختيار الأخير أعلاه، كما يتم مسح هذه الأرقام بعد إنتهاء الاختبار. يمكن تخزين العديد من الاختبارات ومفاتيح الإجابات المختلفة في ملفات مناسبة على القرص المرن الحاوي للبرنامج أو على قرص مرافق، ويتم تخصيص ملف خاص مشفر لكل اختبار ومفتاح إجاباته. ويمكن تنقيح أو تغيير عنوان وتعليمات وفقرات الاختبار ومفتاح إجاباته بعد الإنتهاء من إجراء الاختبار.

٢- إجراء وتصحيح الاختبارات الموضوعية: يمكن استخدام هذا البرنامج، وهو برنامج مرافق لبرنامج B1، في إجراء وتصحيح أي نوع من الاختبارات الموضوعية (اختبارات الصواب والخطأ، اختبارات الاختيارات المتعددة،

اختبارات المطابقة، اختبارات الترتيب، والاختبارات ذات الإجابات القصيرة...
(الخ). وتسمح خيارات البرنامج للمستخدم بالقيام التالي:

⑤ إجراء الاختبار.

⑥ وضع درجات للاختبارات.

⑦ استعراض بنود الاختبار والإجابات.

⑧ طباعة بنود الإجابات ومعدل الدرجات الكلي.

على المستخدم أن يقوم بإدخال كلمة المرور takeit (وهي كلمة المرور التي تأتي مع البرنامج والتي يمكن تغييرها لاحقاً) ورقم الاختبار ورقمه التعريفي. ويعطينا الخيار الأخير أعلاه ورقة مطبوعة تضم رقم الاختبار ورقم المستخدم التعريفي وجدول من ستة أعمدة يدرج العمود الأول رقم السؤال أو الفقرة الاختبارية، ويدرج العمود الثاني أول إجابة يختارها المتحن لكل سؤال (إذ قد يعود المتحن ويغير هذه الإجابة)، أما العمود الثالث فيدرج الزمن بالثواني الذي احتاجه المتحن للتوصل إلى الإجابة، وفي العمود الرابع تُدرج الإجابة النهائية للسؤال (بغض النظر عن عدد المرات التي غير فيها المتحن رأيه واختياره للإجابة). يدرج العمود الخامس الزمن (بالثواني) الذي احتاجه المتحن لاختيار إجابته النهائية، أما العمود السادس فيدرج أحد الحرفين R أو W إشارة لكون الإجابة صحيحة R أو خاطئة W. يمكن للممتحن أن يقوم بإجراء الاختبار واستعراض إجاباته وتصحيح الاختبار، ومن ثم طباعة الإجابات والزمن المستغرق ومجموع درجات الاختبار وذلك بهدف العودة إلى هذه النتائج لاحقاً أو الاحتفاظ بالنتائج كسجل دائم.

٣- التصحيح من أثر التخمين في الاختبارات الموضوعية: يقوم هذا البرنامج القصير بتحويل الدرجات الكلية للاختبارات الموضوعية إلى درجات مصححة

من أثر التخمين. يقوم المستخدم بإدخال عدد الفقرات الكلي للاختبار وعدد الفقرات الصحيحة والخاطئة وعدد الخيارات (البدائل). وتتضمن ورقة النتائج المطبوعة (ملف النتائج) معدل الدرجات قبل التصحيح ومعدل الدرجات المصحح باستخدام معادلة التصحيح من أثر التخمين التقليدية ومعادلة Little ومعادلة Reid (راجع أيكين، ١٩٨٧).

٤- إجراء وتصحيح فقرات المطابقة وإعادة الترتيب: يمكن إجراء فقرات المطابقة أو إعادة الترتيب والتي تتألف كحد أقصى من ١٥-٢٠ جزء باستخدام هذا البرنامج. يتم إدخال أعمدة تضم جذور الأسئلة وأعمدة تضم الإجابات لفقرات المطابقة من خلال الملفين match1 أو match2 أما الإجابات الصحيحة فيتم إدخالها من الملف match3 وعلى الرغم من أنه تم تخزين معطيات عينة لسؤال مطابقة، يمكن للمستخدم تغيير هذه المعطيات بسهولة وإعداد الفقرات التي يريدها. يتم إدخال معطيات فقرات إعادة الترتيب من ملفات rank1 وrank2 وعلى الرغم من أنه تم تخزين معطيات عينة لسؤال إعادة ترتيب، يمكن للمستخدم تغيير هذه المعطيات بسهولة وإعداد فقرات خاصة به. تتألف نتائج فقرات المطابقة من عدد ونسبة المطابقات الصحيحة، أما نتائج فقرات إعادة الترتيب فتتألف من معدلي درجات تم حسابهما بواسطة المعادلات (3.1a) و(3.1b) في الصفحة ٥٩ من أيكين (١٩٨٧a).

٥- وضع درجات ترتيب الفقرات: يقوم هذا البرنامج بإعطاء الدرجات لفقرات الترتيب أو إعادة الترتيب باستخدام طريقة أيكين، ١٩٧٠، يقوم المستخدم بتحديد عدد الفئات وإدراج التصنيفات الصحيحة (المبوبة) وتصنيفات الممتحن لكل فئة. يمكن طباعة الدرجات إما على شكل أعداد عشرية أو أعداد صحيحة مقربة.

٦- وضع الدرجات باستخدام طريقة كاجوري المعدلة: يعطي هذا البرنامج باستخدام تعديل على طريقة كاجوري (أيكين، ١٩٨٣) التوزيع التكراري لدرجات ومجال درجات مجموعة من الطلاب. ويتوجب على المستخدم تحديد عدد الاختبارات أو معدلات الدرجات والحدود القصوى والدنيا للدرجات ووسيط مستوى القدرة للمجموعة (على مقياس بين ١ إلى ١٠٠). راجع أيكين ١٩٨٣ لمعلومات أكثر تفصيلاً عن هذا البرنامج.

٧- احتمالية قيام ممتحنين بأخطاء متماثلة: يستطيع هذا البرنامج اكتشاف الغش في الاختبارات الموضوعية حيث يقوم البرنامج بحساب الإحتمالية ثنائية الحد لقيام ممتحنين بإعطاء عدد محدد من الإجابات الخاطئة المتماثلة لمجموعة من فقرات أسئلة الخيارات المتعددة أو الصواب والخطأ. ويقوم المستخدم في حالة أسئلة الخيارات المتعددة بإدخال عدد الخيارات (البدائل) لكل فقرة وعدد الفقرات التي تمت إجابتها بشكل خاطئ من قبل كلا الممتحنين وعدد الفقرات التي اختار فيها الطالبان الموه نفسه. ويقوم المستخدم، في فقرات الصواب والخطأ (أو الفقرات ذات الخيارين)، بإدخال عدد الفقرات التي أخطأ فيها أحد أو كلا الممتحنين وعدد الفقرات التي أخطأ فيها كلا الممتحنين. راجع أيكين ١٩٩١ لمزيد من المعلومات التفصيلية حول النظرية التي يستند إليها هذا البرنامج.

برامج الفئة C تحليل الفقرات

١- مؤشرات الصعوبة والتمييز للفقرات ذات المرجعية المعيارية: يقوم هذا البرنامج القصير بحساب مؤشر صعوبة فقرات الاختبار (p) ومؤشر تمييز الفقرات (D) لعدد معين من الفقرات. يقوم المستخدم بإدخال العدد الكلي للفقرات وعدد الممتحنين في المجموعات العليا والدنيا (والتي يتم تحديدها

بناءً على المجموع الكلي للدرجات) الذين أجابوا على سؤال ما بشكل صحيح. يتم إدراج قيم P و D تبعاً لرقم الفقرة في النسخة المطبوعة.

٢- مؤشرات الصعوبة والتمييز لل فقرات ذات المرجعية المحكية: يقوم هذا البرنامج القصير بحساب مؤشر صعوبة ومؤشر تمييز الفقرات للاختبارات ذات المرجعية المحكية. وتظهر قيم المؤشرين على شاشة العرض.

٣- تحليل الفقرات الثمانية وإحصائيات الدرجات الكلية واختبارات الدلالة: يقوم هذا البرنامج بحساب وتقييم ثمانية مؤشرات إحصائية لأداء n من الممتحنين في عدد من الفقرات الخيارات المتعددة m . تتضمن المؤشرات قياسات لما يلي:

- ⊗ درجة صعوبة الفقرة.
 - ⊗ الأداء العام للممتحنين في الاختبار.
 - ⊗ التشابه بين إجابات الممتحن على فقرة ما و إجاباته على الفقرات الأخرى.
 - ⊗ التشابه بين إجابات الممتحن على فقرة ما و إجابات ممتحن آخر على نفس الفقرة.
 - ⊗ التشابه في اختيار الموهبات.
 - ⊗ العلاقة بين الفقرة ومعدلات الدرجات الكلية (مؤشر تمييز الفقرة).
 - ⊗ الاختلافات بين معدلات الدرجات الكلية لفقرة ما .
 - ⊗ الاختلافات بين معدلات الدرجات الكلية للممتحنين.
- يتم تعريف المؤشر واختبار الدلالة الموافق له، ويتم تخزين مصفوفة المعطيات ($m \times n$ الممتحن \times البند) في ملف معلومات منفصل. راجع أيكين (١٩٨٩) لمزيد من المعلومات حول البرنامج.

٤- خصائص الفقرات ومنحنى إجابة الفقرات: يقوم هذا البرنامج برسم منحنيات خصائص الفقرات بناءً على المعطيات الأولية أو من خلال نموذج معين للإجابة على الفقرات. يقوم المستخدم بالإشارة إلى ما إذا كان يجب تحديد منحنى خصائص الفقرة من المعطيات الأولية أو من نموذج الإجابة على الفقرات ثم يقوم بإدخال العدد الكلي لمعدلات الدرجات وعدد المجالات ونقطة الوسط لكل مجال وعدد معدلات الدرجات في كل مجال وعدد المتحنيين الذين نجحوا في الفقرة في كل مجال. وفي حال اختيار أحد نماذج الإستجابة للفقرة، يتوجب على المستخدم أن يشير إلى ما إذا يرغب في استخدام نموذج راش أو النموذج ذو المؤشر الواحد أو نموذج ثنائي المؤشرات أو نموذج ثلاثي المؤشرات. يحدد بعد ذلك قيم المؤشرات ويظهر على الشاشة رسم لمنحنى خصائص الفقرة.

٥- نقطة الارتباط ثنائي التسلسل (Point biserial): يقوم هذا البرنامج بحساب معامل نقطة الارتباط ثنائية التسلسل لجملة من معدلات الدرجات المزوجة (التي تم أخذها على شكل أزواج) على متغير ذو حدين (خيارين) أو اختيارات متعددة. يقوم المستخدم بإدخال العدد الكلي للممتحنين ووسطي درجات المحك والانحراف المعياري لدرجات المحك والعدد الكلي للفقرات. كما يقوم المستخدم بالإشارة إلى العدد الكلي للممتحنين الذين نجحوا في فقرة ما ويدخل معدل درجاتهم على مقياس المحك. يتم تخزين معاملات النقطة ثنائية التسلسل تبعاً لترقيم الفقرات في ملف results كما يمكن طباعتها على شاشة العرض أو على الورق.

٦- تحليل فقرات الاختبار الاختيارية: يقوم هذا البرنامج بتقدير مؤشرات الصعوبة والتمييز وثبات مجموعة من فقرات الاختبار الاختيارية (تتألف هذه المجموعة من فقرتين أو أكثر) عندما يطلب من الممتحنين الإجابة على عدد محدد (أكثر من واحد) من الفقرات.

يقوم البرنامج بتحديد ما يلي:

- ⊗ صعوبة الفقرات الاختيارية بناءً على عدد المتحنيين الذين أجابوا عليها
- ⊗ الصعوبة النسبية لأزواج من الفقرات الاختيارية بناءً على المجموعات الفرعية من المتحنيين الذين أجابوا على الفقرتين معاً
- ⊗ الارتباطات النقطية ثنائية التسلسل بين الفقرات الاختيارية التي تمت الإجابة عليها أو لم تتم الإجابة عليها ودرجات المتحنيين في الفقرات المطلوبة
- ⊗ الارتباطات بين درجات الفقرات المطلوبة ودرجات الفقرات الاختيارية
- ⊗ الارتباطات بين أزواج من الفقرات الاختيارية
- ⊗ الحد الأدنى التقديري لثبات الفقرات الاختيارية عندما يكون ثبات الفقرات أو الأقسام المطلوبة أ- معلومة أو ب- غير معلومة
- ⊗ تقديرات الحد الأدنى لثبات الفقرات الاختيارية بناءً على الارتباطات مع أزواج من الفقرات الاختيارية الأخرى عندما يكون ثبات الفقرات الاختيارية غير معلوم. تتألف المعطيات المدخلة من درجات الفقرات المطلوبة الإلزامية والفقرات الاختيارية لكل ممتحن مدرج في بند واحد من الملف المختار. ويتم التعبير عن الفقرات المحذوفة بأعداد أكبر من معدل درجات الأعلى المحتمل لكل فقرة. راجع أيكين (١٩٨٩b) للمزيد من التفاصيل عن الإجراءات الحاسوبية.

٧- تقييم عمل الفقرات التبايني: يقوم هذا البرنامج بحساب قيم دلتا من نسب المتحنيين من مجموعتين والذين نجحوا في فقرات محددة. كما يرسم البرنامج قيم دلتا للمجموعة الثانية بالنسبة لقيم المجموعة الأولى ويولد معادلة الإنحدار الموافقة لها (آنغوف وفورد، ١٩٧٣). كما يُجري البرنامج

اختبار مربعات - كاي لدلالة عمل الفقرات التفاضلي (DIF) باستخدام إجراءات شوينمان (١٩٧٩) وكاميللي (١٩٧٩) ومانتل وهابنتزل (١٩٥٩).

برامج الفئة D: المعايير واختيار العينات:

١- اختيار العينات الطبقي العشوائي: يقوم البرنامج باختيار عينة عشوائية متعددة الطبقات مكونة من m عنصراً من أصل تعداد عام للسكان مكون من n عنصراً. ويتناسب اختيار عدد العناصر من كل طبقة ومن كل بُعد تم تصنيف السكان على أساسه مع التعداد العام للسكان في كل طبقة. وفي حال وجود طبقة واحدة وبعد واحد، يتم اختيار عينة عشوائية بسيطة من مجموع السكان مكونة من عدد العناصر n . أما في حال تساوي عدد العناصر في العينة مع التعداد الكلي للسكان، فيتم إجراء إعادة ترتيب عشوائي لعدد العناصر n .

٢- الخطوط والرتب المتئنية: يقوم هذا البرنامج بحساب التالي (وذلك إنطلاقاً من توزيع تكراري معلوم):

الرتب المتئنية لمعدل درجات معطى أو الخط المتئني الموافق لرتبة متئنية معينة.

٣- الدرجات القياسية: يقوم هذا البرنامج، إنطلاقاً من توزيع تكراري معلوم، بحساب معدلات الدرجات القياسية Z ومعدلات الدرجات المطبّعة $Z_{\text{normalized}}$ ومعدلات الدرجات المشتقة T و Z المكافئة لمعدلات الدرجات تلك.

يقوم المستخدم بتحديد عدد مجالات الدرجات ونقطة الوسط للمجال الأول وعرض المجال، ثم يقوم المستخدم بتحديد تكرار الدرجات على كل من المجالات التي تتبع المجال الأول. وتتألف المخرجات التي تظهر على شاشة العرض من جدول يوضح نقطة الوسط والتكرار والرتبة المتئنية المتوضعة في نقطة الوسط وقيم Z و Z المطبّعة وقيم T لكل مجال.

٤- الاحتمالات والانحرافات الطبيعية: يمكن استخدام هذا البرنامج لحساب إما الإحتمالية الطبيعية لقيمة معلومة للدرجات z أو للدرجات T و $CEEB$ والرتب من تسعة Stanine و NCE (مكافئ المنحنى الطبيعي) الموافقة لاحتمالية تراكمية طبيعية معطاة.

برامج الفئة E: معاملات الثبات

١- معادلات سبيرمان - براون لحساب الثبات: يمكن استخدام هذا البرنامج لتقدير التالي:

• ثبات اختبار تم تطويله: يدخل المستخدم عدد الفقرات الاختبارية للاختبار الأصلي (قبل التطويل) وعدد الفقرات الكلي للاختبار بعد التطويل وثبات الاختبار الأصلي (غير المطول). تظهر قيمة ثبات الاختبار المطول على شاشة العرض.

• عدد الفقرات الاختبارية الإضافية اللازمة لرفع ثبات اختبار ما إلى قيمة محددة: يدخل المستخدم عدد الفقرات الاختبارية للاختبار الأصلي (قبل التطويل) وثبات الاختبار الأصلي (غير المطول) وقيمة الثبات المرغوب فيها للاختبار المطول. ويظهر عدد الفقرات التي يجب إضافتها لرفع قيمة ثبات الاختبار المطول إلى القيمة المرادة على شاشة العرض.

٢- معاملات ثبات كودر - ريتشاردسون: يقوم هذا البرنامج بحساب ثبات الإتساق الداخلي لاختبار ما من خلال معادلات كودر - ريتشاردسون رقم ٢٠ و ٢١، ويقوم المستخدم بإدخال عدد الفقرات الاختبارية والوسط الحسابي لمجموع معدلات درجات الاختبار وتباين مجموع معدلات الدرجات ونسبة عدد المتحنيين الذين أجابوا على الفقرة الاختبارية بشكل صحيح (وذلك لكل فقرة اختبارية على حدة). تظهر معاملات كودر- ريتشاردسون رقم ٢٠ و رقم ٢١ على شاشة العرض.

٣- المعامل ألفا: يقوم البرنامج بحساب المعامل ألفا من خلال المعطيات المدخلة عن طريق لوحة المفاتيح أو عن طريق ملف معطيات، والمعامل ألفا هو مقياس لثبات الإتساق الداخلي لاختبار تمنح فيه الدرجات على أساس النقاط المتعددة لكل فقرة (أي لا يعتمد مبدأ إعطاء نقطة واحدة للإجابة الصحيحة و٠ للإجابة الخاطئة). يقوم المستخدم بإدخال عدد الفقرات الاختبارية وعدد المتحزين والدرجات التي حصل عليها كل فرد في كل فقرة. تظهر القيمة العددية للمعامل ألفا على شاشة العرض.

٤- معامل ثبات المصححين: يقوم هذا البرنامج بحساب معامل الارتباط بين الدرجات التي يضعها عدد من المصححين للورقة الاختبارية ذاتها، كما يحسب البرنامج الارتباط بين مجموع الدرجات العام أو معدلات الدرجات التي توصل إليها أكثر من مصحح.

٥- معامل التوافق: يحدد هذا البرنامج درجة التوافق بين الرتب أو الدرجات التي يمنحها مصححان أو أكثر لورقة اختبارية ما. كما يحسب البرنامج بالإضافة إلى معامل التوافق وقيم مربعات كاي المقابلة والارتباط بين معدلي الرتب والارتباط المتوقع بين مجموعتين من الرتب.

٦- الخطأ المعياري للقياس: يقوم هذا البرنامج بحساب الخطأ المعياري للقياس ومجالات الثقة لـ ٩٠% و ٩٥% و ٩٨% و ٩٩% لمعدل الدرجات الحقيقي لاختبار ما. ولتحديد الخطأ المعياري في القياس ومجال الثقة المقابل له، يقوم المستخدم بإدخال الانحراف المعياري للاختبار ومعامل الثبات.

٧- الثبات والخطأ المعياري لفرق معدلات درجات: يقوم هذا البرنامج بحساب الثبات والخطأ المعياري لفرق معدلات درجات. يدخل المستخدم الانحرافات المعيارية لاختبارين وثباتهما والارتباطات بينهما. عندها يظهر على شاشة

العرض الثبات والخطأ المعياري لفرق معدلات درجات. بعد ذلك، يسأل البرنامج المستخدم ما إذا كان يرغب بحساب مجال الثقة لفرق معدلات الدرجات. وفي حال الجواب بنعم، يطلب من المستخدم الدلالة على النسبة المئوية المطلوبة (٩٠، أو ٩٥، أو ٩٩) لدرجات الاختلاف وللفرق بين معدلي الدرجات. يظهر مجال الثقة لهذا الفرق على الشاشة ويُسأل المستخدم ما إذا كان يرغب بحل مسألة أخرى.

٨- الثبات والنسبة القصوى لثبات مركب اختباري: يقوم البرنامج بحساب الثبات لمجموع أجزاء من اختبار مركب موزونه ويحدد الحد الأقصى للثبات والأوزان المقابلة له. يقوم المستخدم بإدخال عدد المتغيرات للأوزان، ومعامل الثبات والارتباط بين كل زوج من المتغيرات. ويمكن للبرنامج حساب إما: (١) معامل الثبات والخطأ المعياري للقياس المرتبط بأي مجموعة من الأوزان المرجحة للمتغيرات يدخلها المستخدم في البرنامج. و(٢) الحد الأقصى لثبات المركب الاختباري والخطأ المعياري للقياس والأوزان لكل متغير. راجع أيكين (١٩٨٨) للمزيد من المعلومات حول البرنامج.

برامج الفئة F: الاختبارات التحصيلية

١- الاختبارات الحسابية الموضوعية بشكل عشوائي: يقوم هذا البرنامج ببناء وإجراء ووضع درجات اختبار واحد أو أكثر يتألف من عشر فقرات اختبارية وتم وضعه عشوائياً في الجمع أو الطرح أو الضرب أو القسمة أو للرفع على أس للأعداد المؤلفة من خانة واحدة أو خانتين أو خانة واحدة وخانتين في آن واحد.

٢- اختبار التهجي واستخدام الكلمات: يقوم هذا البرنامج بإجراء ووضع درجات اختبار مؤلف من ثلاثة أجزاء:

• الجزء الأول هو اختبار مدى المعرفة للتهجئة الصحيحة لـ ٢٨ كلمة يشيع الخطأ في تهجئتها في الكتابة الرسمية؛

• الجزء الثاني هو اختبار مدى معرفة التهجئة الصحيحة لمجموعة من ٨ كلمات؛

• الجزء الثالث هو اختبار على ٢٠ زوج من الكلمات ذات معانٍ مختلفة.

يعطي البرنامج الدرجات الأولية والنسبة المئوية للإجابات الصحيحة لكل جزء من الاختبار وللإختبار بكامله.

٣- حقائق عن التقدم في السن والحياة الآخرة: يقدم البرنامج اختباراً مكوناً من ٥٠ فقرة من نوع الصواب أو الخطأ حول التقدم في السن والحياة الآخرة. تطبع نسبة الإجابات الصحيحة والدرجات المحتملة المقابلة لها على شاشة العرض.

٤- حقائق عن الموت والاحتضار: يقوم هذا البرنامج بإجراء وإعطاء معدلات الدرجات لاختبار مسأح مؤلف من ٥٠ فقرة صواب أو خطأ حول الموت والاحتضار. تطبع نسبة الإجابات الصحيحة والدرجات المحتملة المقابلة لها على شاشة العرض.

٥- اختبار حول المصطلحات النفسية: يقوم هذا البرنامج باختبار وبشكل شبه شامل عبر ٥٥ فقرة اختبارية ثنائية الخيار معرفة المصطلحات المستخدمة في الدورات التمهيديّة في علم النفس. يسمح البرنامج للممتحن بأن يقوم بمراجعة وتغيير إجاباته على الأسئلة واستعراض الإجابات الصحيحة في نهاية الاختبار. يظهر على شاشة العرض معدل الدرجات الأولية (الرقم الصحيح) ومعدل الدرجات المصحح من أثر التخمين والنسب المئوية لمعدلات الدرجات الأولية ومعدلات الدرجات المصححة من أثر التخمين من معدلات الدرجات المحتملة للاختبار.

٦- اختبار سرعة القراءة ودرجة الاستيعاب للنصوص: لقد تم إعداد هذا الاختبار للتعرف على السرعة ومدى الاستيعاب عند قراءة نص قصير. يطلب من الممتحن قراءة نص ما تبعاً لسرعته في القراءة وضغط أي مفتاح عند الانتهاء. يتم اختبار الممتحن بأربعة أسئلة من نوع الخيارات المتعددة، ويطلب من الممتحن إدخال الرقم المقابل للإجابة الصحيحة. يظهر زمن القراءة (بالثواني) ومعدل الدرجات الأولية ومعدل الدرجات المئوي على شاشة العرض (تم أخذ النص والفقرات الامتحانية من تمارين اختبار GED الرسمي لعام ١٩٩٣ ص ٣٧، وأعيدت طباعتها بموافقة مجلس التعليم الأمريكي).

٧- اختبار سرعة ودقة الطباعة: هذا اختبار لسرعة ودقة طباعة نص قصير. هذا يظهر نص في أعلى شاشة العرض ويطلب من الممتحن طباعته في أسفل الشاشة والضغط على مفتاح الإدخال عند الانتهاء. يظهر على شاشة العرض الزمن (بالثواني) وعدد الأخطاء والنسبة المئوية للإجابات الصحيحة. (تم أخذ النص من تمارين كتيب الوصف العام لاختبار GED 1997-1996 ص ٢٣، أعيدت الطباعة بموافقة مجلس التعليم الأمريكي).

٨- الاختبارات التحصيلية القياسية: يقدم هذا الاختبار البرنامج معلومات تعريفية عامة بالاختبارات التحصيلية المنشورة مبوبة في عشر فئات:

- اختبارات المسح للقراءة.
- اختبارات التشخيص والجاهزية للقراءة.
- اختبارات اللغة الإنجليزية.
- اختبارات اللغات الأجنبية.
- اختبارات الرياضيات.
- اختبارات العلوم.
- اختبارات التاريخ والدراسات الاجتماعية.

⊗ اختبارات مهارات الفن والموسيقى والمكتبات.

⊗ اختبارات الأعمال والاختبارات المهنية.

برامج الفئة G: اختبارات القدرات الحركية والإدراكية - الحسية:

١- اختبار المدى البصري: يقوم هذا البرنامج بقياس درجة الحساسية لأجزاء مختلفة من المجال البصري، وهو اختبار طبي للعيون شائع الاستخدام. يطلب من الممتحن ضبط شدة السطوع على شاشة العرض حتى تصبح رؤية الكلمات أمراً صعباً للغاية ثم يطلب من الممتحن التركيز على بقعة مضيئة صغيرة في وسط الشاشة. ومع كل محاولة من المحاولات الـ ١٠٨ تظهر بقعة أخرى مضيئة في مكان ما من الشاشة وعلى الممتحن أن يضغط على مفتاح المسافة حالما يتمكن من رؤية البقعة المضيئة الثانية. تظهر المحفزات البصرية بشكل عشوائي في ١٢ قطاعاً مختلفاً من الشاشة، ويتم تسجيل إحداثيات المحفز البصري وبعده عن مركز الشاشة ورقم القطاع الذي ظهر فيه والوقت الذي استغرقه الممتحن في الإجابة وذلك لكل محاولة على حدة. وعند الانتهاء من الاختبار يتم حساب متوسط زمن الاستجابة لكل قطاع. ويمكن لتحليل ثنائي الاتجاه (٣-٤) للتباين أن يوضح كيفية تغير زمن الإجابة بتغير موقع المحفز في المجال البصري.

٢- اختبار تمييز اللون: هذا اختبار للقادرة على التمييز بين ألوان أزواج من الوجوه. يطلب من الممتحن الإشارة إلى ما إذا كانت ألوان أزواج الوجوه متماثلة (s) أو مختلفة (d) ويتم تقديم ٢٢ زوجاً من الوجوه (١١) منها ذات ألوان متماثلة و(١١) ذات ألوان مختلفة. تطبع على الشاشة معدلات درجات المجيب بالنسبة لفئات: "أحمر- أخضر" و"أزرق - أصفر" و"الأخطاء المتطابقة" إضافة إلى النسبة المئوية من معدل الدرجات الكلي المحتمل والمتمثل بكل واحد من معدلات الدرجات للفئات الثلاثة المذكورة.

٣- اختبار حدة السمع: يقوم هذا البرنامج بتحديد الحد الأعلى وأعلى تردد للحساسية لشدة الصوت لدى الممتحن. يسمع الممتحن سلاسل مختلفة من النغمات تزداد شدة الصوت فيها بشكل تصاعدي، ويطلب من الممتحن في السلسلة الأولى من النغمات الضغط على مفتاح المسافة عندما يتوقف عن سماع النغمات. أما في السلسلة الثانية فيطلب من الممتحن الضغط على مفتاح المسافة عندما يصل علو النغمة إلى ما يجده حداً أقصى. ومن الجدير بالذكر أن عمر الممتحن وحالته الصحية وغيرها من العوامل تلعب دوراً في الحد المطلق لسماع وتردد الحساسية القصوى لشدة الصوت (يتم حساب التردد بوحددة الهيرتز).

٤- اختبار تمييز الشدة: يتم تقديم ١٤ زوجاً من النغمات كل زوج يتكون من نغمتين لهما إما نفس التردد أو لكل منهما تردد مختلف. ويطلب من الممتحن الضغط على "١" إذا كانت النغمة الأولى أعلى من حيث الشدة والضغط على "٢" إذا كانت الثانية أعلى. تطبع على شاشة العرض عدد الأزواج التي أجاب عليها الممتحن بشكل صحيح والنسبة المئوية لمعدلات الدرجات من أعلى معدل درجات ممكن.

٥- اختبار البحث عن الأرقام: يمثل هذا الاختبار مقياساً لسرعة الملاحظة و الدقة في تحديد المكان والاستجابة لمجموعة من الأرقام. يقوم الممتحن بالبحث في مصفوفة أرقام (١٠ × ١٠) مولدة عشوائياً من الأرقام من (٠) وحتى (٩٩) لتحديد مكان رقم معين وبعد ذلك يقوم بإدخال رقم السطر والعمود اللذين يقع فيهما هذا الرقم. يمكن للقائم على الاختبار (أو الممتحن) أن يقوم بتحديد الحد الأقصى للرقم الواجب البحث عنه والواقع في المجال بين (٠) و(٩٩). وتتص تعليمات الاختبار على ما يلي: انتظر حتى تظهر مصفوفة الأرقام وحتى يتم تعيين الرقم المطلوب ثم أدخل رقم السطر والعمود اللذين يقع فيهما الرقم المطلوب.

٦- اختبار زمن ردّات الفعل دوندرز (أ) و(ب) و(ج): يقوم هذا البرنامج بقياس زمن ردة الفعل البسيطة و المنفردة باستخدام إجراءات دوندرز من الفئات (أ) و(ب) و(ج). وفي الفئة (أ) يقوم الممتحن بالضغط على مفتاح محدد بأسرع وقت ممكن بعد ظهور حرف تحفيزي على شاشة العرض. أما في الفئة (ب) فيظهر واحد من أصل حرفين تحفيزيين على شاشة العرض و على الممتحن الضغط على المفتاح المحدد عند ظهور أحدهما فقط.

أما في الفئة (ج) فيقوم الممتحن بالضغط على مفتاح محدد عند ظهور أحد حرفين تحفيزيين على شاشة العرض وعلى مفتاح آخر عند ظهور الحرف الآخر. تظهر بعد ذلك على شاشة العرض وفي ملف النتائج أزمدة ردة الفعل للفئات (أ) (الزمن الأساسي) و(ب) و(ج) إضافة لزمن التعرف وزمن الاختيار الذي تم حسابه.

٧- اختبار السرعة والصحة للعاملين في الأعمال الكتابية: يتم في هذا الاختبار تقديم ٣٠ زوجاً من الأسماء أو الأرقام بمعدل زوج كل ثانيتين، ويطلب من الممتحن الضغط على المفتاح "s" إذا كان الاسمان أو الرقمان متشابهين وعلى المفتاح "d" إذا كان الاسمان أو الرقمان مختلفين. يتم إعلام الممتحن بأنه لديه دقيقة واحدة للإجابة على الفقرات الثلاثين لذا يجب الإجابة على الفقرات بأسرع ما يمكن. وبعد أن يتم تقديم كافة الفقرات يظهر على الشاشة معدل الدرجات النهائي (الصح و الخطأ) والنسب المئوية المقابلة لها.

برامج الفئة H: اختبارات القدرات الإدراكية

١- الأحاجي (Anagrams): صُمم هذا البرنامج لتحديد الزمن اللازم لتكوين كلمة ذات معنى من كل مجموعة من المجموعات الإثنتي عشر المكونة من خمس أحرف مبعثرة وذلك مع أو بدون أدلة تمهيدية. وفي حالة إعطاء دليل (clue)، يتم إخبار الممتحن بأن الكلمة إما لون أو لها علاقة بالطعام أو العمل،

أما في حالة عدم إعطاء دليل (no-clue) فلا يتم إعطاء أي دليل. تخزن الأحاجي في الملف (scrabble). يتم تخزين عدد المحاولات وزمن الإجابة بالثواني وإجابات الممتحن في الملف results حيث تظهر على الشاشة و يمكن أيضاً طباعتها على الطباعة.

٢- اختبار المدى الرقمي: يقوم هذا البرنامج باختبار القدرة على استدعاء سلسلة من الأرقام المفردة المولدة عشوائياً والمقدمة بمعدل رقم واحد في الثانية والقدرة على تكرارها بنفس الترتيب الذي تم عرضها فيه (digits forward) أو بالترتيب العكسي (digits backward). تنص تعليمات الاختبار على ما يلي وذلك في حالة (digits forward): سوف تظهر سلسلة من الأرقام المفردة من ٢ وحتى ٩ في منتصف الشاشة كل رقم على حدة. حاول أن تحفظ الأرقام من كل سلسلة حتى تتمكن من إعادة طباعتها بنفس الترتيب الذي تم تقديمها فيه. وتنص تعليمات الاختبار في حالة (digits backward) على التالي: سوف تظهر سلسلة من الأرقام المفردة من ٢ و حتى ٩ في منتصف الشاشة كل رقم على حدة، حاول أن تحفظ الأرقام من كل سلسلة حتى تتمكن من إعادة طباعتها بالترتيب المعاكس للترتيب الذي تم تقديمها فيه. وفي كلا الحالتين، يتم إعلام الممتحن بالتعليمات التالية:

بعد أن يتم تقديم كل سلسلة من الأرقام تظهر على الشاشة ومباشرة الرسالة: "Digit؟"، عندها قم بطباعة الرقم التالي ثم قم بالضغط على مفتاح الإدخال. وإذا قام الممتحن بخطأ في أحد السلاسل فيتم تقديم سلسلة جديدة بنفس الطول. يمنح الممتحن درجة واحدة لكل إجابة صحيحة وبعد أن يتم تقديم السلاسل الثمانية لكل حالة يظهر على شاشة العرض معدل الدرجات لتلك الحالة.

٣- سلاسل الأعداد: يقوم هذا البرنامج بتقديم عشر سلاسل مكونة من ست

أرقام و يتوجب على الممتحن تحديد الرقم التالي من السلسلة عبر تطبيق قاعدة حسابية. يتم إعطاء التعليمات التالية:

هذا اختبار لقدرتك على استنتاج الأنماط لمجموعة من الأرقام. سوف يكون هناك عشر محاولات بالمجمل. في كل محاولة سوف تظهر في وسط الشاشة سلسلة من ست أرقام. إن مهمتك ستكون تحديد ماذا يجب أن يكون الرقم التالي (السابع) في السلسلة ثم طباعته والضغط على مفتاح الإدخال. الوقت المتاح هو ٢٠ ثانية لكل سلسلة من السلاسل العشرة. يتم إعلام الممتحن بعد كل محاولة ما إذا كانت إجابته صحيحة أم خاطئة ويتم تخزين رقم المحاولة وزمن الإجابة وكون الإجابة صحيحة أم خاطئة. يمكن طباعة هذه المعلومات بالإضافة إلى معدل الدرجات الكلي للممتحن على الشاشة أو على الورق.

٤- مسائل في البصيرة: يختبر هذا البرنامج القدرة على حل مسائل مكونة من مجموعة من ١٢ كلمة. يتم تخزين هذه المسائل في الملف "probs" والحلول العددية في مفتاح الملف. يطلب من الممتحن قراءة كل مسألة وطباعة و إدخال الإجابة العددية لها (١ أو ٢ أو ٣ إلخ). وفي نهاية الاختبار يظهر على الشاشة معدل الدرجات والنسبة المئوية الصحيحة.

٥- اختباران للقدرات الإدراكية: الاختباران في هذا البرنامج هما مقاييس خاصة للفرص للقدرات الإدراكية والتي لا توجد لها معايير ثابتة. يتكون الاختبار الأول "اختبار المتشابهات اللغوية" من ٢٠ فقرة اختبارية ذات خمس خيارات مصممة لقياس جانب واحد من الذكاء - تعلم العلاقات والارتباطات. أما الاختبار الثاني "اختبار الكلمات المترابطة" فهو اختبار إتمام ٢٠ فقرة اختبارية تم تصميمها تبعاً لنمط اختبار المترابطة البعيدة يقيم القدرات الإبداعية وهو في الواقع اختبار للقدرة اللغوية أو للسهولة في التعامل مع الكلمات. يظهر معدل درجات الممتحن (الرقم الكلي الصحيح) في كل اختبار على شاشة العرض عند انتهاء الاختبار.

٦- مسائل برج هانوي: هذا البرنامج مستوحى من مسألة برج هانوي التقليدية إلا أنه يستخدم الأرقام بدلاً من المكعبات. يطلب من الممتحن تحريك كافة الأعداد في العمود (أ) إلى العمود (ب) على شرط عدم وضع رقم ما فوق آخر أصغر منه وذلك باستخدام عمود (ج) كمرحلة وسيطة أو كمخزن أثناء عملية نقل الأعداد من العمود (أ) إلى العمود (ب). تتم عملية نقل الأرقام من خلال إملء الفراغات في الجملة غير المكتملة: "انقل الرقم العلوي في العمود — إلى العمود —. إطبوع الأحرف أ، ب، ت، في كل فراغ. إضغظ على مفتاح الإدخال بعد القيام بطباعة الحرف. وغالباً ما يبدأ الاختبار بمسألة بسيطة كمسألة مكونة من ثلاثة أرقام و يتطور إلى أربعة أو خمسة أو ستة أرقام ونظرياً حتى عشرة أرقام. إن عدد الحركات في الحل الأمثل هو حيث n هو عدد الأرقام (المكعبات) الواجب تحريكها. ويشار إلى الأخطاء في الحركات عن طريق صوت رنين لتبنيه الممتحن. وبعد التوصل إلى الحل الصحيح يظهر على الشاشة عدد الحركات التي قام بها الممتحن والعدد الأمثل للحركات (العدد الأدنى).

٧- فهرس اختبارات القدرات الإدراكية: يقدم هذا البرنامج معلومات توصيفية عن المئات من اختبارات القدرات العامة والخاصة مقسمة إلى عشر فئات:

- اختبارات الذكاء الفردية
- اختبارات الذكاء الجماعية
- اختبارات القبول المهنية والجامعية
- اختبارات الارتقاء
- اختبارات ذوي الاحتياجات الخاصة
- اختبارات الذاكرة
- الاختبارات العصبية - النفسية.
- اختبارات الاستعدادات المتعددة.
- اختبارات الاستعدادات الخاصة.
- اختبارات المواهب.

obeikandi.com

الملاحق بـ

دور نشر وتوزيع أدوات التقدير النفسية والتعليمية

تحتوي القائمة أدناه على أسماء وعناوين دور النشر والتوزيع الكبيرة ولكن لا يمكن الجزم بكون هذه القائمة كاملة. يمكن الحصول على أسماء وعناوين أخرى في الصفحات ١١٤٥-١١٥٨ من كتاب الاختبارات المنشورة ٤.

Academic Therapy Publications
20 Commercial Boulevard
Novata, CA 94949-6191

ACS DivCHED Examination Institute
Oklahoma State University
107 Physical Sciences
Stillwater, OK 74078

The Alemany Press
P.O. Box 7604
West Trenton, NJ 08628

American College Testing Program
(ACT)
P.O. Box 168
Iowa City, IA 52243

American Council on Education
One Dupont Circle NW, Suite 250
Washington, DC 20036-1163
202-939-9490

American Guidance Service (AGS)
4201 Woodland Road
P.O. Box 99
Circle Pines, MN 55014-1796
800-328-2560

American Orthopsychiatric
Association, Inc.
49 Sheridan Avenue
Albany, NY 12210

Assessment Systems Corporation
2233 University Avenue, Suite 200
St. Paul, MN 55114-1629
612-647-9220

The Australian Council for
Educational Research
19 Prospect Hill Road
Camberwell Vic 3124, Australia
(03) 9277 5656

Ballard & Tighe
480 Atlas Street
Brea, CA 92621

Behavior Data Systems, Ltd.
P.O. Box 44256
Phoenix, AZ 85064-4256
800-231-2401

Behavior Science Systems, Inc.
P.O. Box 580274
Minneapolis, MN 55448
612-929-6220

Center for Applied Linguistics
1118 22nd Street NW
Washington, DC 20037

Consulting Psychologists Press, Inc.
(CPP)
3803 East Bayshore Road
P.O. Box 10096
Palo Alto, CA 94303
800-624-1765

CPPC
4 Conant Square
Brandon, VT 05733
800-433-8234

C.P.S., Inc.
P.O. Box 83
Larchmont, NY 10538
800-433-8324

CTB/McGraw-Hill
20 Ryan Ranch Road
Monterey, CA 93940
800-538-9547

Curriculum Associates, Inc.
5 Esquire Road
North Billerica, MA 01862-2589

Denver Developmental Materials, Inc.
P.O. Box 9619
Denver, CO 80206-9011

DLM Resources
One DLM Park
Allen, TX 75002
800-527-4747

Educational and Industrial Testing
Service (EdITS)
P.O. Box 7234
San Diego, CA 92167
800-416-1666

Educational Developmental Corporation
P.O. Box 470663
Tulsa, OK 74147-0663

Educational Evaluation Enterprises
AWRE
Newnham, Gloucestershire GL14 1ET
England

Educational Testing Service (ETS)
ETS Test Collection (30-B)
Rosedale Road
Princeton, NJ 08540

Educators Publishing Service, Inc.
31 Smith Place
Cambridge, MA 02138-1100

GED Testing Service
One Dupont Circle NW
Washington, DC 20036

Grune & Stratton, Inc.
465 South Lincoln Drive
Troy, MO 63379

Hawthorne Educational Services, Inc.
800 Gray Oak Drive
Columbia, MO 65201
800-542-1673

Hilson Research, Inc.
P.O. Box 150239
Kew Gardens, NY 11415-0239
800-926-2258

Hodder & Stoughton Education
Headline PLC, 338 Euston Road
London NW1 3BH
England

Human Sciences Research Council
P.O. Box 32410, 1
Braamfontein 2017
South Africa

Industrial Psychology International, Ltd.
4106 Firestone Road
Champaign, IL 61821
217-398-1437

Institute for Personality and Ability
Testing (IPAT)
P.O. Box 1188
Champaign, IL 61824-1188
800-225-IPAT

Psychologists and Educators, Inc.
P.O. Box 513
Chesterfield, MO 63006

Psychometric Affiliates
Box 807
Murfreesboro, TN 37133-0807
615-898-2565

Publishers Test Service, CTB/McGraw-Hill
20 Ryan Ranch Road
Monterey, CA 93940
800-538-9547

Ramsey Corporation, Boyce Station
Offices
1050 Boyce Road
Pittsburgh, PA 15241

Reitan Neuropsychology Laboratories
2920 South 4th Avenue
Tucson, AZ 85713-4819
602-882-2022

Richardson, Bellows, Henry & Co., Inc.
1140 Connecticut Avenue NW
Washington, DC 20036

Riverside Publishing
425 Spring Lake Drive
Itasca, IL 60143-2079
800-323-9540

Scholastic Testing Service, Inc. (STS)
480 Meyer Road
P.O. Box 1056
Bensenville, IL 60106-1617
800-642-6STS

Science Research Associates (SRA)
155 North Wacker Drive
Chicago, IL 60606

Selby MillSmith Ltd.
30 Circus Mews
Bath BA1 2PW
United Kingdom
01225 446655

Sigma Assessment Systems, Inc.
P.O. Box 610984
Port Huron, MI 48061-0984
800-265-1285

Slosson Educational Publications, Inc.
P.O. Box 280
East Aurora, NY 14052-0280
800-828-4800

SOI Systems
P.O. Box D
45755 Goodpasture Road
Vida, OR 97488
541-896-3936

Special Child Publications
P.O. Box 33548
Seattle, WA 98133

SRA/London House
9701 Higgins Road
Rosemont, IL 60018
800-221-8378

Stoelting
620 Wheat Lane
Wood Dale, IL 60191
630-860-9700

U.S. Employment Service, Division of
Program Planning and Operations
U.S. Department of Labor
601 D Street NW
Washington, DC 20213

U.S. Military Entrance Processing
Command Testing Directorate
2500 Green Bay Road
North Chicago, IL 60064

Western Psychological Services (WPS)
12031 Wilshire Boulevard
Los Angeles, CA 90025-1251
800-648-8857

Wide Range, Inc.
P.O. Box 3410
Wilmington, DE 19804-0250
800-221-WRAT

Wonderlic Personnel Test, Inc.
1509 North Milwaukee Avenue
Libertyville, IL 60048-1380
800-963-7542

Jastak Associates, Inc.
P.O. Box 3410
Wilmington, DE 19804-0250
800-221-WRAT

Jossey-Bass
615 Montgomery Street
San Francisco, CA 94111

Lafayette Instrument Company
P.O. Box 5729
3700 Sagamore Parkway North
Lafayette, IN 47903

Law School Admission Council
11 Dupont Circle NW, Suite 150
Washington, DC 20036

MacMillan Education
Houndsmills
Basingstoke, Hampshire RG21 2XS
England

Martin M. Bruce
50 Larchwood Road
Box 248
Larchmont, NY 10538
914-834-1555

McGraw-Hill/London House
9701 West Higgins Road
Rosemont, IL 60018-4720
800-237-7685

Merrill (Charles E.) Publishing Company
1300 Alum Creek Drive
Box 508
Columbus, OH 43216
800-848-1567

Metritech, Inc.
4106 Fieldstone Road
P.O. Box 6489
Champaign, IL 61826-6479
217-398-4868

MHS
908 Niagara Falls Boulevard
North Tonawanda, NY 14120-2060

Moreno Educational Co.
7050 Belle Glade Lane
San Diego, CA 92119

National Council on Economic Education
1140 Avenue of the Americas
New York, NY 10036

NCS Assessments
P.O. Box 1416
Minneapolis, MN 55440
800-627-7271

Nelson Canada
1120 Birchmount Road
Scarborough Ontario M1K 5G4
Canada

New Zealand Council for Educational
Research
Education House West
178-182 Willis Street
Box 3237
Wellington, New Zealand

NFER-Nelson Publishing Co., Ltd.
Darville House
2 Oxford Road East
Windsor, Berkshire SL4 1DF
England

Personnel Press
191 Spring Street
Lexington, MA 02173

PRO.ED
8700 Shoal Creek Boulevard
Austin, TX 78757-6897
512-451-3246

Psychological and Educational
Publications, Inc.
1477 Rollins Road
Burlingame, CA 94010-2316
800-523-5775

Psychological Assessment Resources, Inc.
(PAR)
P.O. Box 998
Odessa, FL 33556
800-331-TEST

Psychological Corporation (The)
555 Academic Court
San Antonio, TX 78204-2498
800-211-8378

Psychological Publications, Inc.
P.O. Box 3577
Thousand Oaks, CA 91359-0577
800-345-TEST

Psychological Test Specialists
Box 9229
Missoula, MT 59807

المحقق

الموزعون التجاريون لمنتجات الحاسب وخدمات التقدير النفسي والتعليمي

American College Testing
2201 North Dodge Street
Iowa City, IA 52243
319-337-1000

American Guidance Service
4201 Woodland Road
P.O. Box 99
Circle Pines, MN 55014-1796
800-328-2560

Assessment Systems Corporation
2233 University Avenue, Suite 200
St. Paul, MN 55114-1629
612-647-9220

Behavior Data Systems, Ltd.
P.O. Box 32938
Phoenix, AZ 85064-2938
800-231-2401

Caldwell Report
1545 Sawtelle Boulevard, Suite 14
Los Angeles, CA 90025
310-476-3133

Consulting Psychologists Press, Inc.
3803 East Bayshore Road
P.O. Box 10096
Palo Alto, CA 94303
800-624-1765

Educational and Industrial Testing
Service
P.O. Box 7234
San Diego, CA 92167
619-222-1666

Hilson Research, Inc.
P.O. Box 150239
Kew Gardens, NY 11415-0239
800-926-2258

Institute for Personality & Ability
Testing, Inc.
P.O. Box 1188
Champaign, IL 61824-1188
800-225-4728

Integrated Professional Systems, Inc.
5211 Mahoning Avenue, Suite 135
Youngstown, OH 44515
216-799-3282

Century Diagnostics, Inc.
2101 East Broadway, Suite 22
Tempe, AZ 85282
602-966-6006

CFKR Career Materials, Inc.
11860 Kemper Road, Unit 7
Auburn, CA 95603
916-880-2357

Clinical Psychometric Research
P.O. Box 619
Riderwood, MD 21139
301-321-6165

Computerized Psychological Diagnostics,
Inc.
1101 Dove Street, Suite 225
Newport Beach, CA 92660
714-833-7931

Conover Company Ltd.
1050 Witzel Avenue
Oshkosh, WI 54901
414-231-4667

NFER-Nelson Publishing Company Ltd.
2 Oxford Road East
Windsor, Berkshire, SL4 1DF
England

Planet Press
P.O. Box 3477
Newport Beach, CA 92663-3418
714-650-5135

PRO.ED
8700 Shoal Creek Boulevard
Austin, TX 78757-6897
512-451-3246

Psychological Assessment Resources, Inc.
P.O. Box 998
Odessa, FL 33556
800-331-TEST

Psychological Corporation (The)
555 Academic Court
San Antonio, TX 78204-2498
800-228-0752

Psychological Psoftware Company
12486 Brickellia
San Diego, CA 92129
619-484-8877

Jastak Associates, Inc.
P.O. Box 3410
Wilmington, DE 19804-0250
800-221-WRAT

Life Science Associates
One Fenimore Road
Bayport, NY 11705
516-472-2111

Metritech, Inc.
4106 Fieldstone Road
P.O. Box 6489
Champaign, IL 61826-6479
800-747-4868

Multi-Health Systems, Inc.
908 Niagara Falls Boulevard
North Tonawanda, NY 14120-2060
800-456-3003

NCS Assessments
5605 Green Circle Drive
P.O. Box 1416
Minneapolis, MN 55443
800-NCS-7271

Slosson Educational Publications, Inc.
P.O. Box 280
East Aurora, NY 14052-0280
800-828-4800

StatSoft
2300 East 14th Street
Tulsa, OK 74104-4442
918-583-4149

Test Agency
Cournswood House, North Dean
High Wycombe, Bucks HP14 4NW
England
44-24-024-3384

University Associates
8517 Production Avenue
San Diego, CA 92121-2280
619-578-5900

Psychologistics, Inc.
P.O. Box 033896
Indialantic, FL 32903
407-259-7811

Risk & Needs Assessment, Inc.
P.O. Box 32818
Phoenix, AZ 85064-2818
800-231-2401

Riverside Publishing Company (The)
8420 Bryn Mawr Avenue
Chicago, IL 60631
800-767-TEST

Selby MillSmith
30 Circus Mews
Bath BA1 2PJ
Avon, England
(225) 446655

Sigma Assessment Systems, Inc.
P.O. Box 610984
Port Huron, MI 48061-0984
800-265-1285

Vocational Research Institute
1528 Walnut Street, Suite 1502
Philadelphia, PA 19102
800-874-5387

Western Psychological Services
12031 Wilshire Boulevard
Los Angeles, CA 90025-1251
800-648-8857

Wonderlic Personnel Test, Inc.
1509 North Milwaukee Avenue
Libertyville, IL 60048-1380
800-963-7542

قائمة المصطلحات والفردات

اختبار القدرة: هو اختبار يقيس درجة قدرة الفرد على أداء مهام معينة أو احتراف مهنة معينة.

الاستعداد الأكاديمي: ويطلق عليه أيضاً اسم الاستعداد المدرسي، وهو القدرة على التعامل بنجاح مع المهام المدرسية أو ما يشابهها. تختبر معظم اختبارات الذكاء الاستعداد الأكاديمي عند الأفراد.

التحصيل: وهو درجة نجاح أو إنجاز فردٍ ما في مجال أو مشروع معين؛ وكذلك الدرجة المعطاة لاختبار للتحصيل.

السلوك التكيفي: وهو درجة قدرة الفرد على التواصل الفعال والمناسب مع البيئة المحيطة به.

الاختبار التكيفي: وهو إجراء اختباري غالباً ما يتم عن طريق الحاسب وتتنوع فيه الأسئلة والمهام تبعاً للقدرة المتوقعة للممتحن أو تبعاً لصفات الممتحن الخاصة أو لإجاباته على الأسئلة السابقة.

التأثير المضاد (السلبي): وهو كل ما له تأثير سلبي على قبول بعض الأفراد المنتمين إلى جماعات معينة في بعض الوظائف والمؤسسات التعليمية.

المقاييس العاطفية: وهي اختبارات أو أدوات اختبار مصممة لقياس المزاج والشاعرية والاهتمامات والمواقف وأنماط الشخصية وغيرها من أنواع السلوك العاطفي للأفراد.

الدرجات المكافئة للعمر: راجع المعيار العمري.

المعيار العمري: وهو وسيط درجات اختبار تحصيلي أو اختبار للاستعداد أُجري لأطفال من عمر معين.

ثبات الصورة المكافئة: وهو مؤشر للثبات (معامل التكافؤ) يُحدد عن طريق الربط

بين درجات الفرد التي حصل عليها في صورة معينة للاختبار مع الدرجات التي حصل عليها في صورة أخرى من الاختبار ذاته.

التقييم البديل: وهي طريقة غير تقليدية للتقييم تعتمد على الطلب من الممتحن أن ينتج أو يولد استجابات إبداعية لفقرات الاختبار بدلاً من الإجابة على فقرات محدودة الإجابات.

اختبار التشابهات: وهو اختبار يتطلب من الممتحنين تحديد العلاقات إما من حيث التشابه أو الاختلاف بين شيئين أو أكثر. مثال: الأحمر بالنسبة للورود هو كماذا بالنسبة للبنفسج: ١- الأزرق ٢- الأخضر ٣- البرتقالي ٤- الأصفر.

طريقة وضع الدرجات التحليلية: وهي إجراء لوضع الدرجات في الاختبارات المقالية يتم فيه تقييم إجابات الممتحنين بالنظر إلى عاملي الأسلوب والمحتوى.

اختبار الفقرات المثبتة (anchor test): وهي عدد من فقرات الاختبار تبقى ثابتة في صور متعددة من ذات الاختبار وتساعد في معادلة درجات كافة صور الاختبار.

الاستعداد: وهو القدرة على تعلم كيفية تأدية مهمة ما أو مهارة معينة. وقد كان المعتقد السائد في الماضي هو أن الاستعداد قدرة تولد مع الفرد ولا تُكتسب بالخبرة والتدريب.

اختبار الاستعداد: وهو مقياس للقدرة على الاستفادة من تدريب أو خبرة إضافية تمكن الشخص من أن يصبح أكثر براعة أو خبرة. وتفيد الفقرات المختلفة في اختبارات الاستعداد في التنبؤ بالأداء المستقبلي للأفراد في مجالات العمل والتعليم.

التقدير: وهو تقييم وجود أو حجم صفات شخصية معينة. ويتضمن تقدير assessment السلوك البشري والعمليات العقلية عدداً من الإجراءات مثل الملاحظة والمقابلات وسلالم التقدير وقوائم الرصد ومقاييس السمات وتقنيات التنبؤ والاختبارات.

مقياس الصوت: وهو أداة لقياس الحدة السمعية يقوم على تقديم أنغام صافية تتعدد درجات عمقها وذبذبتها ضمن المجال السمعي العادي. يتم اختبار السمع في الأذنين، ثم تستخدم النتائج في رسم مخطط بياني سمعي يحدد حدة الممتحن السمعية بالنسبة لكل ذبذبة وكل أذن.

التقدير الواقعي: وهو تقدير assessment الأداء بالاعتماد على مهام حقيقية ومستوحاة من الواقع.

سعة الطيف: وهو مصطلح أطلقه ل. ج. كرونباخ على مجموعة المحكات التي يمكن التنبؤ بها من خلال اختبارٍ ما، وكلما تعددت المحكات التي يمكن لاختبارٍ ما التنبؤ بها اتسع طيف هذا الاختبار. راجع مصطلح الأمانة.

العمر القاعدي: وهو أعلى مستوى عمري في اختبارات الذكاء (كالطبقات الأخيرة من اختبار ستانفورد - بينيه). يتجاوز الممتحن وبنجاح جميع الاختبارات الفرعية الأدنى منه أو المطابقة لمستواه.

النسبة القاعدية: وهي نسبة الأشخاص ضمن تعداد سكاني معين يمثلون خاصية، حالة أو سلوك معينين.

التحيز: وهو الاتجاه في العديد من المقاييس السياكومترية إلى التوصل لنتائج أعلى أو أدنى من النتائج الفعلية؛ كذلك التحيز مع فئة من الناس على حساب فئة أخرى. راجع مصطلح العدالة.

السطف العمري: وهو أدنى مستوى عمري في اختبارات الذكاء (كالطبقات الأخيرة من اختبار ستانفورد - بينيه). يفشل الممتحن في جميع اختباره الفرعية.

التصنيف: وهو استخدام درجات الاختبار لوضع الأفراد ضمن مجموعات مختلفة.

تقنية كلوز: وهي إجراء اختباري يتم فيه مسح عدد من الكلمات من نص ما ثم يطلب من الممتحن ملء فراغات النص. وتقييم قدرة الممتحن على القراءة من خلال قدرته على فهم النص وملء الفراغات بالشكل المطلوب.

اختيار العينات: وهو إجراء لأخذ العينات يتم فيه تقسيم السكان إلى مجموعات أو أقسام. ويتناسب عدد الوحدات المختارة عشوائياً من كل مجموعة مع إجمالي عدد الوحدات في هذه المجموعة.

التدريب والإعداد: وهو إجراء تعليمي قصير الأمد يهدف إلى تحسين درجات الممتحنين في الاختبارات التي سيخضعون لها في المستقبل. يتم تدريب الممتحنين على

الاختبار مسبقاً من خلال التعرف على نمط الأسئلة والتدرب على استراتيجيات الاختبار.

المعامل ألفا: وهو معامل ثبات الإتساق الداخلي وهو مناسب للاستخدام في الاختبارات ذات الأسئلة التي تمنح درجة ١ أو ٠ والأسئلة ذات الدرجات المتعددة؛ كذلك معامل الارتباط المتوقع لاختبار ما مع نظيره يحتوي عدداً مماثلاً من الفقرات أو الأسئلة.

معامل المطابقة Concordance: وهو معامل العلاقة بين الدرجات التي يمنحها مصحح ما لورقة اختبارية ما والدرجات التي يمنحها مصحح آخر لنفس الورقة.

معامل التكافؤ: وهو معامل للثبات (إرتباط) يمكن الحصول عليه عن طريق إجراء شكلين مختلفين من الاختبار ذاته للأفراد ذاتهم. راجع ثبات الصور المتكافئة.

معامل الإتساق الداخلي: وهو معامل للثبات مبني على تقديرات الإتساق الداخلي للاختبار؛ معاملات التجزئة النصفية، معاملات كودر - ريتشاردسون ومعامل ألفا.

معامل الإستقرار: وهو معامل للثبات يمكن الحصول عليه عن طريق إجراء الاختبار ذاته للأفراد ذاتهم في توقيتين مختلفين. راجع ثبات إعادة الاختبار.

معامل الإستقرار والتكافؤ: وهو معادل للثبات يمكن الحصول عليه عن طريق إجراء شكلين مختلفين من الاختبار ذاته للأفراد ذاتهم في توقيتين مختلفين.

الإدراك: وهو كل ما له علاقة بالفكر من ذاكرة وتفكير وحل المشكلات وغيرها.

التقدير الإدراكي: وهو مقياس للعمليات العقلية كالإدراك والذاكرة والتفكير والحكم والإستدلال.

النمط الإدراكي: وهو اتجاه أو نمط الإدراك والتذكر والتفكير الذي يحدده الفرد لنفسه بغرض فهم العالم من حوله والتأقلم معه (مثال: الاعتمادية والاستقلال، التعقل والاندفاعية، التحكم بالذات داخلياً أو تحكم العالم الخارجي بالفرد).

التشارك Communalilty: وهو نسبة التباين في متغير تم قياسه يمكن أن تعزى إلى التباين الذي يشترك بها هذا المتغير مع غيره من المتغيرات.

الدرجات المركبة: وهي المجموع المباشر أو المثلث لدرجات اختبارين أو أكثر أو قسمين من نفس الاختبار.

الاختبار التكيفي المحوسب: وهو إجراء اختبائي عن طريق الحاسب لتكيف فيه الأسئلة والمهام اللاحقة في الاختبار تبعاً لإجابات المتحن على الأسئلة السابقة.

ثبات المطابقة: يقوم عدد من المصححين بإعطاء درجة عددية معينة لدرجة وجود سلوك أو خاصية ما عند عينة كبيرة من الناس، ثم يتم حساب معامل المطابقة (وهو مؤشر للتوافق بين آراء المصححين).

الصدق التلازمي: وهو درجة ارتباط الدرجات التي يحصل عليها مجموعة من الناس في اختبار سايكومتري معين مع الدرجات التي يمكن التوصل إليها عن طريق مقياس آخر للخاصية نفسها التي يحاول الاختبار قياسها.

مجال الثقة: وهو المجال الذي تقع ضمنه درجة المتحن الحقيقية في اختبار ما أو في متغير المحك. ويتراوح مجال الثقة فيبين ٩٥% و٩٩%. راجع الخطأ المعياري للتقدير والخطأ المعياري للقياس.

وزن الثقة: وهي طريقة موضوعية لوضع الدرجات يعتمد فيها الثقل العددي المعطى لإجابة صحيحة في اختبار ما على درجة ثقة المتحن في صحة إجابته.

صدق البناء: وهي درجة ارتباط الدرجات التي يحصل عليها الفرد في اختبار سايكومتري مصمم لقياس خصائص معينة مع المقاييس التي تقيس السلوك في مواقف تحدد فيه الخاصية موضع القياس سلوك الفرد.

الفقرات الإنشائية: وهي أسئلة أو مسائل تتطلب إجابتها كتابة أو رسم بياني أو صورة أو غيرها مما يتطلب الإنشاء. ويعرف هذا النوع من الأسئلة أيضاً باسم "الأسئلة المفتوحة".

صدق المحتوى: وهو درجة تأكيد المختصين والخبراء على أن الاختبار أو المقياس السايكومتري يقيس في الواقع ما كان قد صمم لقياسه.

التصحيح من أثر التخمين: وهي معادلة تطبق على الدرجات الخام للاختبارات لتصحيح آثار الإجابات التي يخمنها المتحن. ومن أكثر معادلات التصحيح هي تلك التي تطرح جزءاً من عدد الإجابات الصحيحة من عدد الإجابات الخاطئة.

الإرتباط: وهي درجة العلاقة بين متغيرين كالعلاقة بين الاختبار ومقياس المحك. **معامل الارتباط:** وهو مؤشر عددي لدرجة العلاقة بين متغيرين ويتراوح عادة بين $-1,00$ (علاقة سلبية كاملة) و $0,00$ (لا علاقة أبداً) إلى $+1,00$ (علاقة إيجابية كاملة). وهناك نوعان من معامل الارتباط وهما المعامل المولد للعزوم ومعامل النقطة ثنائية التسلسل).

المحك: وهو مقياس أو متغير تُقاس أو تُقيّم درجات أي اختبار سايكومتري بالإستناد إليه. ويُحدد صدق أي اختبار سايكومتري يُستخدم في إختيار أو تصنيف الأشخاص عن طريق قدرة هذا الاختبار على التنبؤ بمحك سلوكي معين مرتبط بالموقع الذي يتم إختيار الأشخاص له أو تصنيفهم ضمنه.

تشوه المحك: وهو تأثير أي عامل كان على المحك يؤدي إلى أن يصبح المحك غير صالح لقياس إنجازات الفرد. فعلى سبيل المثال تُستخدم نتائج اختبارات الاستعداد للتنبؤ بدرجات التحصيل المدرسية ولكن استخدام المدرسين لنتائج اختبارات الاستعداد لتحديد درجات الطالب يجعل هذه الدرجات محكاً غير صالح لإثبات صدق اختبار الاستعداد وبالتالي فقد تم تشوه المحك.

الاختبار ذو المرجعية المحكية: وهو اختبار يُقيد محتواه بمواصفات خاصة ويخدم عدداً معيناً من الأغراض المحددة. وغالباً ما يكون الهدف من الاختبار هو تحديد درجة تحصيل الممتحن بالمقارنة مع أهداف تعليمية معينة. راجع الاختبارات ذات المرجعية المعيارية.

الصدق المرتبط بالمحك: وهو درجة قدرة اختبار ما على قياس ما يريد هذا الاختبار أن يقيسه. ويمكن تحديد هذا الأمر بالنظر إلى الإرتباط بين درجات الاختبار ومقياس المحك لسلوك ما.

السلوك الحاسم Critical incident: وهو أي سلوك يتوقف عليه الأداء الفعال لفردٍ ما في وظيفة ما (مثال: "ينظف مكان العمل قبل المغادرة" و"يتعامل مع الزبائن بلباقة").

الصدق عبر المجموعات Cross validation: وهي طريقة لتحديد صدق اختبار ما مع مجموعة جديدة من الناس بعد أن تم إثبات صدق هذا الاختبار في التنبؤ بمحك

سلوكي ما عند جماعة بشرية أخرى. يتم هذا عن طريق إجراء هذا الاختبار للمجموعة الجديدة. وغالباً ما يتقلص معامل الصدق مع المجموعة الثانية من المتقدمين إذ أن عامل الحظ يضغط معامل الصدق في المرة الأولى للاختبار.

القدرة المتبلورة (g): وهو مصطلح أطلقه ر. ب. كاتل على القدرات العقلية (القدرات المعرفية) المكتسبة من الخبرة والتعليم.

الاختبار العادل حضارياً: وهو اختبار يُفترض في مواده أن تكون متوفرة لجميع الشعوب والحضارات فلا يعاقب هذا الاختبار أية مجموعة حضارية بسبب قلة خبرتها في أمور معينة. وحتى الآن لم تكمل الجهود لتصميم اختبارات عادلة حضارياً بالنجاح. الدرجات الفاصلة (الحد الأدنى للنجاح): يتم قبول جميع المتقدمين الذين حصلوا على درجات أكثر أو مساوية للدرجات الفاصلة ورفض المتقدمين الذين حصلوا على درجات دون هذا الحد. وتعتمد الدرجات الفاصلة للاختبار على صدق الاختبار، النسبة المحددة مسبقاً لعدد الناجحين بالنسبة لعدد المتقدمين، وغيرها من العوامل.

الدرجات المشتقة: وهو معدل للدرجات يتم التوصل إليه عن طريق إجراء عملية حسابية بسيطة على الدرجات الأولية للاختبار مثل ضرب معدل الدرجات الأولية في عدد ثابت أو/و إضافة هذا العدد الثابت للدرجات. راجع الدرجات المعيارية، درجات T، ودرجات z.

العمر التطوري (الارتقائي) (DA): وتمثله درجة التحصيل في جداول جيزل الإرتقائية.

معدل التطور (DQ) وهو مؤشر ($DQ = 100 \times DA/CA$) يماثل بشكل تقريبي العمر العقلي للرضع ويلخص سلوك الرضيع تبعاً لتقييم جداول جيزل.

معدل الذكاء الإنحرافي: وهو معدل لدرجات اختبار الذكاء يُستخلص عن طريق تحويل معدل الدرجات الخام لاختبار ذكاء إلى توزيع درجات وسطه الحسابي ١٠٠ وإنحرافه المعياري ثابت (فهو ١٦ لاختبار ستانفورد - بينيهو ١٥ لاختبارات ويكسلر).

الاختبار التشخيصي: وهو اختبار تحصيلي مؤلف من عدد من المهارات التي تشكل مادة دراسية معينة يهدف إلى تحديد نقاط القوة والضعف لفرد ما في المجال

موضع الاختبار. ويتوفر هذا النوع من الاختبارات في مجالات القراءة والحساب والتهجئة.

عمل الفقرات التبايني (DIF): يحدث هذا عندما تتغير مؤشرات فقرات الاختبار عند تغير المجموعات موضع الاختبار (مثال: أن تجد إحدى المجموعات فقرات امتحانية ما أسهل من غيرها أو أكثر تحيزاً من غيرها).

صدق التمييز: وهي حالات وجود علاقات إرتباط ضعيفة بين أداة قياس سايكومترية ما ومقاييس أخرى للمفهوم العقلي.

المموه distracter: وهي اجابة خاطئة تماماً توضع ضمن إختيارات الإجابة للأسئلة المتعددة الإختيارات.

التقييم الديناميكي: وهو أسلوب في التقييم يعتمد على نموذج درّس- اختبر - درّس ويتم فيه اختبار الفرد (الاختبار الأولي) ثم تدريبه على مواد الاختبار ثم إعادة اختبار (الاختبار اللاحق). ويشكل اختلاف الأداء بين الاختبارين مقياساً لقابلية التعلم عند الفرد. راجع منطقة التطور الممكن.

مرض صعوبة الحساب (ديسكالكوليا): وهي عدم القدرة على إجراء العمليات الحسابية.

مرض صعوبة القراءة والكتابة (ديسليكسيا): وهو خلل في القدرة على القراءة يرتبط مع خلل القدرة على تفسير العلاقات المكانية أو الجمع بين المعلومات البصرية والسمعية.

المتخلفين عقلياً الذين يمكن تعليمهم (EMR): وهم الأطفال الذين تم تشخيص نسبة بسيطة من التخلف العقلي لديهم (وهم أولئك الذين تتراوح معدلات ذكائهم بين ٥١ و ٦٩). ويمكن لهؤلاء الأطفال أن يصلوا إلى مستوى تعليمي بين الصنفين الثالث والسادس، كما يمكنهم أن يقوموا بعمليات حسابية بسيطة بالإضافة إلى القراءة والكتابة. راجع المتخلفين عقلياً الممكن تدريبهم.

الهدف التعليمي: وهو هدف أو برنامج تعليمي أو تأهيلي يحدد ما يجب على المتعلم أن يقدر على فعله مع نهاية الدرس أو الدورة التعليمية.

التصحيح عن طريق الاستبعاد: وهو إجراء لوضع الدرجات يقوم فيه المصحح بالإشارة إلى الإجابات الخاطئة لفقرة اختبارية ما بدلاً من تصحيح الإجابة الأفضل للفقرة فقط.

طريقة المؤشر (الخط) المئيني المعادل: وهي طريقة تقليدية لتحويل وحدات درجات إختبار ما إلى وحدات درجات لاختبار مكافئ. يتم تحويل درجات كلا الاختبارين إلى رتب مئينية وتوضع النتائج في جدول يتم التوصل إليه عن طريق معادلة الدرجات الواقعة على الخط المئيني p للاختبار الأول مع الدرجات الواقعة على الخط المئيني p للاختبار الثاني.

الصور الامتحانية المكافئة: راجع الشكل الامتحاني الموازي.

الاختبار المقالي: وهو نوع من الاختبارات يطلب فيه من المتحنيين الإجابة على عدد من الأسئلة بجواب واحد مطول. ويتم تقييم الإجابات من قبل المدرس أو المقيم بشكل ذاتي. راجع الاختبار الموضوعي.

جدول التوقع: وهو جدول يوضح تكرار أو نسبة المتحنيين الواقعين ضمن فئة معينة (حزمة درجات) من متغيرٍ تنبؤي (اختبار) يتوقع أن يقعوا ضمن فئة معينة (فئات درجات) من متغيرٍ المحك.

الصدق المظهري: وهي درجة كون مظهر أو محتوى الاختبار ومواده (أو فقراته أو ما شابه ذلك) تعطي الانطباع بأن الاختبار أو المقياس السايكومتري يقيس ما يجب أن يقيسه.

العامل factor: وهو بُعد أو صفة أو خاصية للشخصية يمكن الحصول عليها عن طريق تحليل عوامل مصفوفة الارتباط التي تم حسابها من درجات عدد كبير من الأفراد في اختبارات مختلفة أو فقرات من اختبارات مختلفة.

التحليل العاملي: وهو إجراء حسابي لتحليل مصفوفة الارتباط بين المقاييس لتحديد العوامل (أو المفاهيم) القادرة على تفسير الارتباط.

أحمال العامل Factors loading: هو، في التحليل العاملي، الارتباطات الناتجة (الأوزان) بين الاختبارات (أو غيرها من المتغيرات) والعوامل المستخلصة.

دوران العامل: Factor rotation وهو إجراء حسابي يتم إجراؤه على مصفوفة المعاملات بهدف تبسيطها لتفسير نتائجها. ويتم هذا عن طريق زيادة عدد تحميلات العاملين الأعلى والأدنى في المصفوفة. ويمكن أن يكون دوران العامل عامودياً بحيث تشكل المعاملات الناتجة زوايا قائمة مع بعضها، أو منحرفاً بحيث تشكل محاور العامل الناتج زوايا حادة أو منفرجة.

العدالة: وهي درجة عدم تحيز الدرجات النهائية في الاختبارات، أو كون هذه الدرجات تتبأ بشكل متعادل بالأداء على المحك لمجموعات مختلفة من الأفراد.

الخطأ السلبي غير الحقيقي: وهو خطأ في الإختيار أو خطأ في القرار التشخيصي يتبأ فيه الإجراء التقديري بنتيجة سيئة (مثال: تحصيل دون المستوى، أداء متدني، الاضطراب العقلي).

الخطأ الإيجابي غير الحقيقي: وهو خطأ في الإختيار أو خطأ في القرار التشخيصي يتبأ فيه الإجراء التقديري بنتيجة جيدة (مثال: تحصيل أعلى من المستوى، أداء جيد، غياب الاضطراب العقلي).

الأمانة: وهو محدودية سعة طيف الاختبار أو غيره من أدوات القياس. ويتبأ الاختبار ذو سعة الطيف المحدودة بشكل جيد بعدد محدد من المحكات. راجع سعة الطيف أو النطاق.

السؤال ذو البدائل المحدودة: وهو سؤال متعدد الاختيارات يتألف من جذر وعدة إجابات محتملة.

الاختبار ذو المستويات المرنة Fexilevel test: وهو اختبار يتألف من عدة أسئلة (فقرات) مرتبة من الأسهل إلى الأصعب ويبدأ الممتحن الاختبار من المنتصف وعندما يعطي الممتحن إجابة خاطئة عن سؤال ما يُطرح عليه سؤال أسهل، أما عندما يعطي إجابة صحيحة فيُطرح عليه سؤال أصعب.

القدرة الضطرية (g_f): وهي مصطلح أطلقه ر. ب. كاتل على القدرة العقلية الكامنة داخل الأفراد والتي تحددها الوراثة مثل قدرات حل المشاكل والاستجابات الإبداعية (غير التقليدية) للأمور.

التقدير التكويني: وهو طريقة لتقييم الأداء يتم عن طريقها إما تطوير أساليب التعليم والتدريس أو التعرف على نقاط الضعف والقوة لدى الطلاب بهدف تطوير مواد تعليمية تثرى معرفة المتفوقين أو تحل المشاكل التعليمية للطلاب الضعيفين. راجع التقييم التراكمي.

قاعدة الأربعة أخماس: وهي قاعدة للاختيار تقضي بأن تدني نسبة معدل اختيار أي عرق أو جنس أو نوع في أي إجراء للاختيار دون نسبة الأربعة أخماس (٨٠٪) من العرق أو الجنس الذي له نسبة الاختيار الأعلى له آثار سلبية، وبالتالي فإجراء الاختيار هذا غير قانوني.

العامل g: وهو معامل عام للذكاء إفترضه سبيرمان لتفسير الترابط الشديد بين اختبارات الذكاء المختلفة.

معامل التعميم: وهو معامل عددي يعد مؤشراً على درجة القدرة على التعميم بالانتقال من العينة موضع الاختبار إلى السكان بشكل عام.

نظرية التعميم: وهي نظرية لدرجات الاختبارات وللتشكيل الإحصائي المرتبط بها والذي يعطي تصوراً لدرجات اختبار ما كعينة من ضمن مجموعة كبيرة من العينات. يستخدم تحليل إجراءات التباين، لتحديد القدرة على تعميم الدرجات كقيم عامة (كونية) لأداء الممتحنين وال فقرات الامتحانية والظروف الامتحانية.

"القاعدة الذهبية": وهي حل وسط توصلت إليه شركة تأمين القاعدة الذهبية وخدمة الاختبارات التعليمية (ETS) وافقت فيه ETS على إجراء اختبار لندوبي التأمين مؤلف من فقرات اختبارية لا تحتوي على فروقات كبيرة بين العرق الأبيض والعرق الأسود.

معيار الصف Grade norm: وهو وسيط درجات الاختبار لعدة أطفال منتمين إلى سنة دراسية واحدة.

استراتيجية التلقين التدريجي Graduated Prompting Strategy: وهي استراتيجية للتقييم الديناميكي يقوم فيها مُجري الاختبار بتقديم عدد من التلميحات السلوكية بغرض تعليم الممتحنين القواعد اللازمة لاجتياز الاختبار. وغالباً ما تكون التلميحات الموجودة في نص معد سابقاً والتي لا تعتمد على إجابات الممتحنين عامة جداً أول الأمر ثم تصبح أكثر تحديداً.

الاختبار الجماعي: وهو اختبار يجريه فرد واحد لعدد من الممتحنين في نفس الوقت.

التصحيح الشمولي: وهو إجراء لوضع الدرجات (وغالباً ما يستخدم في الاختبارات ذات الأسئلة المقالية) يتم فيه تقييم أداء الممتحن بشكل عام ووضع درجات لإجاباته بناءً على ذلك دون إعطاء درجات مختلفة لعناصر الإجابة المختلفة. راجع وضع التصحيح التحليلي.

الصدق المتزايد: وهو زيادة في صدق الاختبار تنتج عن استخدام اختبار جديد في عملية الإنتقاء وتتجاوز هذه الزيادة صدق كل إجراءات الاختيار السابقة.

الاختبار الافرادي Individual test: وهو اختبار يُجرى لكل فرد على حدة.

الموافقة المطلعة: وهي الموافقة على اختبار نفسي أو المشاركة في دراسة نفسية بعد الإطلاع على جميع الحقائق أو المحاذير المتعلقة بالاختبار.

الذكاء: تتعدد تعريفات الذكاء حيث يرى أ. بينيه أنه القدرة على الحكم بشكل جيد والفهم بشكل جيد والمحاكمة بشكل جيد، ويعرفه ل. م. تيرمان بالقدرة على التفكير المجرد، إلا أن اختبارات الذكاء بشكل عام تقيس قدرة الأفراد على النجاح في المهام المدرسية وشبهاتها.

معدل الذكاء (IQ): وهو معدل درجات مشتق، استخدم أصلاً في وضع درجات على مقياس ستانفورد - بينيه للذكاء. يتم حساب معدل الذكاء النسبي بتقسيم عمر الممتحن العقلي على عمره الفعلي ثم يضرب الناتج ب ١٠٠، بينما يتم حساب معدل الذكاء الانحرافي بضرب الناتج Z الموافق للدرجات الأولية لاختبار ذكاء ما بالعدد الذي يمثل الانحراف المعياري لمعدلات الذكاء الانحرافية ثم إضافة ١٠٠ للناتج.

اختبار الذكاء: وهو اختبار نفسي مصمم لقياس درجة استعداد الفرد للعمل المدرسي وغيره من النشاطات التي تدخل فيها القدرة اللغوية وقدرة حل المشكلات.

الفقرات الاختبارية المتداخلة: وهي فقرات اختبارية تعتمد أو تتأثر فيها الإجابة على أحد الأسئلة على إجابات الأسئلة الأخرى.

الاتساق الداخلي: وهو درجة كون جميع الفقرات الامتحانية قادرة على قياس المفهوم ذاته أو المتغير ذاته. وإن ثبات الاختبار الذي يقاس بواسطة معادلات سبيرمان - براون أو كودر - ريتشاردسون أو كرونباخ - ألفا هي مؤشر على الإتساق الداخلي للاختبار.

ثبات المصححين: وهي نوع من الثبات تحدها درجات للاختبار يضعها مصححان لاختبارات عينة من الأفراد.

المؤشر الفئوي: وهو مقياس يشير فيه تساوي الفروقات العددية إلى تساوي الإختلافات في الصفة أو الخاصية موضع القياس. وتعد مقاييس الحرارة (درجة مئوية وفهرنهايت) ومقاييس الدرجات القياسية مقاييس فئوية.

الفقرة الاختبارية: وهي إحدى الوحدات أو الأسئلة أو المهام التي تؤلف الاختبار السايكومتري.

تحليل الفقرات: وهو مصطلح عام يطلق على إجراء تقدير ثبات أو فائدة مجموعة ما من الفقرات الامتحانية.

التحيز في فقرات الاختبار Itembias: وهو درجة كون فقرة امتحانية ما تقيس مفاهيم مختلفة عند مجموعات عرقية أو حضارية أو إقليمية أو جنسية مختلفة.

منحنى خصائص الفقرة: وهو رسم بياني يستخدم في تحليل الفقرات الاختبارية يتم فيه رسم عدد المتحنيين الذين اجتازوا فقرة ما بنجاح إلى جانب درجات الاختبار الكلية.

مؤشر صعوبة الفقرة: وهو مؤشر يدل على صعوبة أو سهولة فقرة اختبارية ما بالنسبة إلى عدد من المتحنيين. ويتم قياس صعوبة فقرة ما عن طريق حساب النسبة المئوية (p) لعدد المتحنيين الذين اختاروا الإجابة الصحيحة لهذه الفقرة.

مؤشر تمييز الفقرة: وهو مقياس لفعالية فقرة اختبارية ما في التمييز بين الأفراد الذين حصلوا على درجات عالية في الاختبار (أو في متغير آخر للمحك) وأولئك الذين حصلوا على درجات متدنية.

تقاطع الفقرات: وهو الدرجة التي تقيس فيها فقرتين مختلفتين المتغير ذاته مما يؤدي إلى ارتباط درجات الفقرتين.

منحنى الإجابة على الفقرة: وهو رسم بياني يوضح عدد المتحنيين الذين أجابوا على الفقرة بشكل صحيح يُرسم إلى جانب محك داخلي (مجموع درجات الاختبار) أو خارجي للأداء.

نظرية الإستجابة للفقرة IRT: وهي نظرية ل فقرات الاختبار يتم فيها التعبير عن درجات الاختبارات على شكل درجات تقديرية ضمن مجال القدرة الكامنة.

الحصول على عينات الفقرات: وهو إجراء لإنتقاء فقرات لاختبارات فرعية من ضمن بنك للفقرات. وتقدم عينات مختلفة من الفقرات لمجموعات مختلفة من المتحنيين.

تحليل الوظيفة: وهو مصطلح عام للإجراءات التي تحدد المهام المحددة التي تتطلبها عليها وظيفة ما. ومن الضروري أن يتم هذا الإجراء قبل بناء اختبار للتبؤ بالأداء الوظيفي للأفراد.

معادلات كودر - ريتشاردسون: وهي معادلات تستخدم في حساب ثبات الإتساق الداخلي لإجراء اختبار يُصحح على مبدأ $1 - \rho$ (درجة واحدة للإجابة الصحيحة و٠ للإجابة الخاطئة) مرة واحدة.

اختبار اللغة: وهو اختبار مؤلف من فقرات لفظية وحسابية، أي فقرات تتطلب استخدام لغة ما (الأحرف أو الأرقام).

نظرية السمات الكامنة Latent trait theory: وهي واحدة من عدد من النظريات (كنظرية منحنى خصائص الفقرة ونموذج راش) مع الإجراءات الإحصائية المرافقة لها والتي تربط بين درجات الفقرات ودرجات الاختبارات من جهة وتقدير معين لموقع الفرد ضمن السمات أو المجال الافتراضي للقدرة الكامنة من جهة أخرى؛ تُستخدم هذه النظرية في تحليل الفقرات وقياس الاختبارات.

الصعوبات التعليمية: وهي صعوبات يواجهها أفراد معدلات ذكائهم حول المعدل أو أعلى من المعدل في القراءة والحساب والكتابة والتهجئة وغيرها من المهارات الأكاديمية.

النتع الخالي من التجريح Least Stigmatizing label: وهي تصنيف أو وصف لحالة تم تشخيصها يعبر عن الحالة بشكل دقيق دون أن يجرح الأفراد المعنين أو أن يسبب لهم وصمة اجتماعية.

تحليل الإنحدار الخطي: وهو إجراء لتحديد المعادلة الجبرية للخط الأنسب للتنبؤ بقيمة متغير مستقل ما من متغير تابع أو أكثر.

الربط: وهو إجراء من نظرية IRT لمعادلة اختبارين عن طريق تحويل مؤشرات الفقرات لاختبار ما إلى مؤشرات للفقرات لصورة ثانية من نفس الاختبار الأمر الذي يضع المؤشرات المتوافقة للاختبارين على المقياس العددي ذاته.

المعايير المحلية: الرتب المئوية والدرجات القياسية وغيرها من المعايير الموافقة لدرجات اختبار أولية لمجموعة صغيرة من الممتحنين المحليين.

اختبار البراعة: راجع الاختبار المبني على المحكات.

فقرات المطابقة: وهي نوع من فقرات الاختبار تطلب من الممتحنين المطابقة بين جذور السؤال الموجودة في عمود أو قائمة ما والإجابات الموجودة في قائمة (عمود) آخر.

القياس: وهو إجراء لحساب كمية أو مقدار مفهوم ما أو أي كيان آخر؛ كذلك تعيين أرقام للأحداث والأشياء.

الوساطة Mediation: وهي استراتيجية للتقدير الديناميكي يتفاعل فيها الممتحن مع القائم على الاختبار بشكل متواصل وذلك بهدف زيادة احتمالية توصل الممتحن إلى الحل المناسب لمسألة اختبارية ما.

العمر العقلي (MA): وهو معدل درجات مشتق لاختبار ذكاء كاختبار ستانفورد - بينيه. يتوافق العمر العقلي لممتحن ما مع العمر الفعلي (الزمني) لعينة من الأطفال تمثل المجتمع عمرها الفعلي واحد ومعدل درجاتها مساوٍ لدرجات هذا الممتحن. راجع معدل الذكاء النسبي (IQ).

الوضع الدراسي المبني على العمر العقلي: وهو مؤشر للسنة الدراسية التي يعتبر فيها طالب ما فعال عقلياً.

الموهوبون عقلياً: أي شخص أداؤه الفكري أعلى بكثير من المعدل وغالباً ما يكون ذكاؤه بين ١٣٠ و ١٤٠ أو أعلى.

المعاقون عقلياً: أي شخص أداؤه الفكري أدنى بكثير من المعدل وغالباً ما يكون ذكاؤه بين ٧٠ و ٧٥ أو أدنى.

المتغير الوسيط: وهو متغير يؤثر على الارتباط بين متغيرين آخرين.

الاختبار متعدد المستويات: وهو اختبار تم تصميمه ليناسب عدداً من المستويات والأعمار؛ كذلك اختبار منفصل يتم إعداده لكل مستوى على حدة.

بطارية الاستعدادات المتعددة: وهي بطارية اختبارات معايرة Conormed تم تصميمها لتقدير القدرات الفعلية.

فقرات الخيارات المتعددة: وهي فقرات اختبارية تتألف من جذر السؤال (تصريح أو سؤال أو عبارة أو غيرها) وعدد من الاختيارات أو البدائل واحد منها فقط صحيح.

معامل الارتباط المتعدد (R): وهو مقياس لدرجة الارتباط العام بين عدد من المتغيرات لها محك واحد للمتغير. يتراوح R بين $+1$ و -1 لعدة متغيرات واحد منها بعد متغيراً محكياً، ويتراوح الارتباط المتعدد العام لمجموعة من اختبارات الاستعداد مع معدلات الدرجات المدرسية بين ٦٠، و ٧٠، وهي نسبة ارتباط معتدلة.

الدرجات الفاصلة المتعددة Multiple cutoff: طريقة لإنتقاء الموظفين أو الطلاب تعتمد على إحراز الممتحن لعدد من الحدود الدنيا للدرجات في عدد من محكات الإنتقاء ليتم قبوله في الوظيفة أو المؤسسة التعليمية.

تحليل الإتحاد المتعدد: وهو إجراء إحصائي لتحليل مساهمة متغيرين مستقلين أو أكثر في التنبؤ بقيمة متغير تابع.

المعايير الوطنية National norms: الرُتب المئينية والدرجات القياسية وغيرها من المعايير المبنية على عينة على مستوى الدولة. راجع: المعايير المحلية والمعايير.

التقدير العصبي - النفسي: وهو قياس الأداء الإدراكي والإدراكي - الحسي والحركي لتحديد موضع ودرجة وآثار الخلل الوظيفي العصبي والإصابات العصبية.

الإضطراب العصبي - النفسي: وهو اضطراب في الجهاز العصبي المركزي ترافقه أعراض نفسية.

المقياس الإسمي: وهو أدنى درجات القياس، وتستخدم فيه الأرقام كصفات أو أسماء للأشياء لا تعبر عن ترتيب أو كمية معينة.

الاختبار الالفاظي Nonverbal test: وهو اختبار لا يتطلب من الممتحن استخدام الكلمات المحكية أو المكتوبة بل استخدام الإشارة أو هز الرأس أو البناء أو التعامل مع الأشياء أو الإستجابة لأسئلة الاختبار بطريقة لفظية.

التوزيع الطبيعي: وهو توزيع سلس للدرجات يأخذ شكل الجرس. وهو متناظر حول الوسط الحسابي وتعبّر عنه دالة رياضية دقيقة. وغالباً ما تتوزع درجات مجموعة كبيرة من الممتحنين بشكل طبيعي.

الدرجات القياسية المطبّعة Normalized standard scores: وهي درجات قياسية (z) يتم الحصول عليها بتحويل درجات الاختبارات الأولية بطريقة تقود إلى توزيع هذه الدرجات توزيعاً طبيعياً.

مجموعة المعيار Norm group: عينة من الأفراد تم قياس الاختبار عليها.

الاختبار ذو المرجعية المعيارية: وهو اختبار يتم تفسير نتائجه بالإستناد إلى معايير تم الحصول عليها من عينة من الناس تمثل المجتمع. راجع الاختبار ذو المرجعية المحكية.

المعايير Norms: هي قائمة تضم معدلات درجات مع الرتب المثوية والدرجات القياسية وغيرها من الدرجات المشتقة لعينة قياس الاختبار الموافقة لهذه الدرجات.

الاختبار الموضوعي: وهو اختبار يتم تصحيحه (وضع درجاته) عن طريق مقارنة إجابات الممتحنين مع مفتاح للإجابات تم إعداده مسبقاً، على عكس الاختبارات الذاتية. ومن الأمثلة على الاختبارات الموضوعية اختبارات الخيارات المتعددة واختبارات الصواب والخطأ.

الدوران المائل Oblique rotation: في التحليل العاملي، هو الدوران الذي تُشكل محاور العامل مع بعضها زوايا حادة أو منفرجة. ينتج عن ذلك الارتباط بين العوامل.

ثبات الفقرات الفردية والزوجية: الارتباط بين مجموع درجات الفقرات ذات الأرقام الزوجية ومجموع الدرجات الفقرات ذات الأرقام الفردية والذي يتم تصحيحه باستخدام معادلة سبيرمان - براون للثبات. راجع معادلة سبيرمان - براون.

كتيب الاختبار: وهو اختبار يتألف من عدد من الفقرات تقيس نواح متعددة من النشاط العقلي. اختبار أوتيس - لينون للقدرة المدرسية واختبار هينمون - نلسون

للقدرة العقلية هي اختبارات تقدم على شكل كتيب اختبائي. راجع كتيب الاختبار اللولبي.

الفقرات المفتوحة الإجابة: راجع الفقرات الإنشائية.

التعريف العملياتي: وهو تعريف يحدد ما يجب على فرد ما أن يعمل له ليحس أو يجرب المفهوم المراد تعريفه.

العملياتيية Operationalism: وهي عقيدة فلسفية تتطوي على التعبير عن مفهوم ما أو مصطلح علمي ما أو نظرية ما من خلال العمليات أو الوظائف التي تعرف أو تجسد هذا المفهوم أو المصطلح أو النظرية.

اختبارات المهارات اللغوية الشفهية (Oracy Test)^(١): وهو اختبار لمهارات الإستماع والكلام.

الاختبار التحصيلي الشفهي: وهو اختبار يجيب فيه الممتحنون شفهيًا على أسئلة كتابية أو شفوية.

المقياس الرتبي: وهو نوع من أنواع القياس تعبر فيه أرقام الأشياء عن رتب أو مراتب الأحداث والأشياء (كترتيب الإنتهاء في مسابقة جري أو غيرها).

الدوران العمودي Orthogonal rotation: في التحليل العاملي، وهو تدوير يحافظ على استقلالية العامل أي يُبقي على الزوايا القائمة بين العوامل مما يعني عدم وجود أي ارتباط بين هذه العوامل.

الاختبار الخارج عن المستوى: وهو اختبار يناسب عمر أو سنة دراسية أدنى أو أعلى من عمر الممتحن أو سنته الدراسية.

الصور المتكافئة: وهما اختباران تتشابه فقراتهما من حيث درجة الصعوبة وترتفع نسبة الإرتباط بينهما. كما تتقارب الدرجات التي يحصل عليها الأفراد في الاختبارين بشكل كبير.

(١) كلمة oracy هي كلمة جديدة في اللغة الإنجليزية لم تتم إضافتها إلى معظم المعاجم حتى الآن وهي كلمة مركبة على نمط كلمة literacy التي تشير إلى مهارات الكتابة والقراءة. ومن الملاحظ أن هذه الكلمة تجمع بين كلمتي oral ولاحقة cy - (الترجم).

ثبات الصور المكافئة: وهو مؤشر لثبات الاختبار يُحدد عن طريق حساب الارتباط بين درجات صورتين متشابهتين من الاختبار ذاته.

تحليل المسار: وهو إجراء إحصائي يتم فيه تحليل العلاقات بين المتغيرات لتحديد تسلسلها أو تتابعها السببي.

الخط المئيني: وهو معدل لدرجات اختبار ما تقع النسبة المئوية لمعدلات درجات معينة تحته أو عليه.

الحزمة المئينية: وهي عدد من الرتب المئينية تشير لإحتماليات محددة إلى وقوع الدرجات الفعلية لمتحن ما ضمنها.

المعايير المئينية: وهي قائمة بالدرجات الخام والنسب المئوية الموافقة لها الناتجة عن عينة القياس والتي تقع تحت الخط المئيني المعطى.

الرتب المئينية: وهي النسب المئوية للدرجات التي تقع تحت معدل درجات معين في التوزيع التكراري للدرجات؛ كذلك النسبة المئوية الموافقة لمعدل درجات معين.

التقدير الأدائي: وهو نوع من التقييم يعتمد على البناء أو الإبداع أو العرض والتوضيح. وتتوفر في معظم الحالات عدة طرق للتقدير وأكثر من طريقة واحدة للإجابة بشكل صحيح.

اختبار الأداء: وهو اختبار يطلب من الممتحنين التعامل مع عدد من الأشياء الملموسة؛ تعتبر هذه الاختبارات نقيض اختبارات الورقة والقلم (أو الاختبارات الموضوعية).

معامل نقطة الارتباط ثنائي التسلسل Point-biserial coefficient: وهو معامل للارتباط بين متغير يحتمل احتمالين أو قيمتين فقط (dichotomous) ومتغير مستمر يحتمل عدد من الاحتمالات (continuous)^(٢).

مقياس النقاط Point scale: وهو اختبار يتم فيه إعطاء درجات لإجابات الطلاب (٠ أو ١ أو ٢... إلخ) بناءً على صحة الإجابة وشموليتها.

(٢) بعض المتغيرات كالجنس لا يحتمل إلا متغيرين (ذكر أو أنثى) أما العمر أو معدل الذكاء فهي أمثلة على المتغيرات المتعددة الاحتمالات، (المترجم).

ملف إنجاز الطالب (البورتفوليو): وهو مجموعة من أعمال الطالب الأدائية أو الإنتاجية يتم تجميعها خلال فترة زمنية معينة بغرض تقييمها.

اختبار القوة: وهو اختبار يُعطى فيه الممتحنون وقتاً طويلاً للإجابة على أسئلة الاختبار التي قد تكون على درجة عالية من الصعوبة.

الصدق التنبؤي: وهي درجة قدرة درجات اختبار ما على التنبؤ بالأداء في محك ما يتم قياسه لاحقاً؛ يتم التعبير عن الصدق التنبؤي على شكل ارتباط بين الاختبار (متغير التنبؤ) ومتغير المحك.

خصوصية التواصل Privileged communication: وهي أي تبادل للمعلومات بين الفرد ومحاميه أو طبيبه أو القسيس أو زوجه/زوجته يُمنح فيه مُعطي المعلومات الحق في منع الطرف الثاني من الإدلاء بشهادة رسمية تكشف أي من هذه المعلومات.

الاختبار التشخيصي: وهو اختبار يستخدم للتنبؤ بقدرة فرد ما في مادة معينة (كاختبار الجاهزية للقراءة).

المهارات الحركية - النفسية: وهي المهارات التي تتطلب حركات معينة كرمي الكرة والإمساك بها وإدخال شيء في شيء آخر وغيرها من الحركات التي تسمح للفرد بالتعامل مع الأشياء (المهارات الرياضية).

المعايرة العرقية Race norming: وضع درجات معيارية للاختبارات تستند إلى عينة من عرق أو فئة إثنية معينة، وتقييم المتقدمين بناءً على العرق الذي ينتمون إليه.

العينة العشوائية: وهي عينة من الملاحظات (كدرجات الاختبارات) يتم الحصول عليها من عينة من المجتمع تم اختيارها بطريقة تضمن لكل فرد من العينة الفرصة نفسها في الاختيار.

المدى (المجال): وهو مقياس خام لحدود أو تنوع مجموعة من درجات الاختبار يتم حسابه عن طريق طرح معدل الدرجات الأدنى من معدل الدرجات الأعلى.

التواصل: وهو علاقة ودودة وحميمة تنشأ بين الممتحن والقائم على الاختبار.

نموذج راش: وهو نموذج أحادي المؤشر لحساب صعوبة الفقرات يستخدم في تحليل فقرات الاختبار وقياس الاختبارات. يستند هذا النموذج إلى فرضية أن مؤشرات

التخمين وتمييز الفقرة هي مؤشرات لا تؤخذ بعين الاعتبار نظراً لصغر قيمها. ويعتمد نموذج راش، شأنه كشأن باقي نماذج السمات الكامنة، على تحديد موقع الطالب في المجال الافتراضي للقدرة الكامنة بناءً على أدائه في فقرات الاختبار المختلفة.

معدل الذكاء النسبي Ratio IQ: وهو معدل للذكاء يتم الحصول عليه بتقسيم درجات العمر العقلي لفرد ما في أحد اختبارات الذكاء (كالنسخ الأقدم من اختبار ستانفورد - بينيه) على عمره الفعلي ثم ضرب الناتج بالعدد ١٠٠، راجع معدل الذكاء الإنحرافي.

المقياس النسبي: وهو مؤشر للقياس قيمته الدنيا هي الصفر وتدل النسب العددية المتساوية بين فواصله على نسب متساوية للصفة أو الخاصية موضع القياس. وعلى حين لا يتم عادةً قياس المتغيرات النفسية باستخدام المؤشر النسبي، يتم استخدامه لقياس متغيرات أخرى كالطول والعرض والوزن وغيرها من المتغيرات الفيزيائية.

اختبار الجاهزية: وهو اختبار يقيس مهارات ومعرفة فرد ما لتحديد جاهزيتها لتعلم المواد الدراسية المعقدة كالقراءة والرياضيات والتهجئة.

فقرات إعادة الترتيب: وهي نوع من الفقرات الاختبارية التي تتطلب من الممتحن إعادة ترتيب أجزاء السؤال وفقاً لتسلسلها الصحيح.

معادلة الإنحدار: وهي معادلة خطية للتنبؤ بدرجات المحك من درجات متغير تنبؤي واحد أو أكثر؛ كذلك إجراء يستخدم في عمليات الإنتقاء (للطلاب والموظفين) أو في حساب المخاطر أو في عمليات التشخيص.

الإنحدار نحو الوسط الحسابي: وهي نزعة لدى درجات الاختبارات وغيرها من المقاييس النفسية في الاقتراب من الوسط الحسابي عند إعادة الاختبار؛ ويزداد اقتراب معدل درجات ما من الوسط الحسابي عند إعادة الاختبار عندما يكون معدل الدرجات الأصلي شديد التطرف.

الثبات: الحد الذي بالنسبة له تقيس أداة تقدير نفسي أي شيء باتساق. وغالباً ما تتميز الدرجات الناتجة عن المقاييس الثابتة بأنها قريبة من حيث القيمة العددية من الدرجات الحقيقية للممتحنين وخالية نسبياً من الأخطاء.

معامل الثبات: وهو مؤشر عددي بين ٠,٠٠ و ١,٠٠ يعبر عن ثبات أدوات القياس. يتم حساب ثبات الاختبارات باستخدام عدة طرق كطريقة إعادة الاختبار والصور المتكافئة والإتساق الداخلي.

قواعد التقييم (rubrics): وهو إجراء أو مخطط يتألف من عدد من المعايير يتم تقييم أداء المتحنيين على أساسها.

اختبار الاستعداد المدرسي: وهو اختبار يتنبأ بقدرات الأفراد على تلقي المعلومات التي تُدرّس على مستوى المدرسة، ويشابه هذا النوع من الاختبارات اختبارات الذكاء التقليدية.

المسح: هو مصطلح عام يستعمل لأي عملية انتقاء ويكون عادة غير دقيق يتم فيه الاحتفاظ ببعض المتقدمين والإستغناء عن الآخرين.

الاختبار الآمن: وهو اختبار يتم ضمن إجراءات أمنية مشددة يضمن عدم اختبار أي فرد غير الأفراد المعنيين وعدم نقل أوراق الإجابة خارج قاعة الإمتحان حتى من قبل القائمين على الاختبار.

الإنتقاء: وهو عملية استخدام عدد من الاختبارات لاختيار المتقدمين الذين ترتفع احتمالية نجاحهم في المقرر أو التخصص أو المنصب الذي تقدموا له. ويتم قبول الأفراد الذين تقع معدلات درجاتهم فوق أو على الحد الفاصل للدرجات، بينما يتم رفض كل من تقع درجاته دون هذا الحد.

نسبة (معدل) الانتقاء: وهو نسبة المتقدمين الذين تم انتقائهم للإلتحاق بمقرر أو برنامج تدريبي أو عمل ما.

المعالجة التلاحقية (التتابعية) للمعلومات: عملية عقلية تتم فيها معالجة المعلومات بشكل متتابع أو متلاحق. ومن الأمثلة على المعالجة التلاحقية محاولة تذكر سلسلة من الأرقام. راجع المعالجة المتزامنة.

الاختبار التتابعي: وهو نوع من الاختبارات تعتمد فيه الأسئلة التي تعطى للممتحن على إجاباته السابقة. يطلق على هذا النوع من الاختبارات كذلك اسم الاختبارات التكيفية أو الاختبارات المكيفة.

الفقرات ذات الإجابات القصيرة: وهي فقرات لا تتطلب إلا إجابة قصيرة لملء الفراغ أو الإجابة على السؤال.

المعالجة المتزامنة للمعلومات: عملية عقلية يتم فيها التأليف بين أجزاء متفرقة من المعلومات في آن واحد. راجع المعالجة التلاحقية.

مخطط سنيلين: وهو لوحة مؤلفة من أحرف صغيرة وكبيرة تقيس دقة وحدة البصر عند الأفراد.

معادلة سبيرمان - براون: وهي معادلة عامة لتقدير ثبات اختبار ما (r_{11}) تم زيادة فقراته بمقدار m . ففي المعادلة ، $r_{mm} = mr_{11}/[1+r_{11}(m-1)]$ ، حيث r_{11} هي ثبات الاختبار الأصلي (قبل الزيادة) و m هي العدد المضاف.

الأطفال ذوي الخصوصية (الإحتياجات الخاصة): وهم الأطفال الذين يعانون من مشاكل جسدية أو نفسية أو إدراكية تجعل عملية وصولهم إلى أهدافهم وتلبية حاجاتهم الخاصة أصعب مما هي عليه مع غيرهم من الأطفال.

أدوات التحديد: وهي ألفاظ مثل: كافة ودائماً وإطلاقاً فقط تشير إلى أن سؤال صواب أو خطأ ما هو خطأ على الأرجح، أو ألفاظ مثل: غالباً وأحياناً وعادةً تشير إلى أن سؤال صواب أو خطأ ما هو صحيح على الأرجح.

الإعاقات التعليمية الخاصة: راجع الإعاقات التعليمية.

الخصوصية: وهو الجزء من تباين الاختبار الذي يُعزى إلى عوامل خاصة بهذا الاختبار.

اختبار السرعة: وهو اختبار مؤلف من عدد كبير جداً من الأسئلة السهلة نسبياً ووقته محدود جداً مما يؤدي إلى فشل معظم المتحنيين في إتمام الإجابة على كافة الأسئلة ضمن المدة الزمنية المتاحة. وغالباً ما تكون اختبارات المهارات الميكانيكية والمكتبية والحركية - النفسية اختبارات سرعة.

اختبار الكتيب اللولبي: وهو نوع من الاختبارات يتم فيه ترتيب الفقرات من الأسهل إلى الأصعب على شكل لولبي مما يعني أن الفقرات المتشابهة تتوزع على مراحل الاختبار المختلفة وتأتي مع الفقرات المماثلة في الصعوبة وليس في المحتوى.

معامل التجزئة النصفية: وهو طريقة لحساب ثبات الاختبار تعتمد على تطبيق معادلة سبيرمان - براون ($m = 2$) على الارتباط بين نصفي الاختبار ذاته (كالفقرات ذات الأرقام الزوجية والفقرات ذات الأرقام الفردية).

الإنحراف المعياري: وهو الجذر التربيعي للتباين؛ يستخدم لقياس إنتشار مجموعة من معدلات الدرجات.

الخطأ المعياري للتقدير: وهو مقدار الخطأ في تقدير درجات فرد ما في متغير المحك بالإستناد إلى درجاته في متغير التنبؤ.

الخطأ المعياري للقياس: وهو تقدير للإنحراف المعياري للتوزيع الطبيعي للدرجات التي يُفترض أن يحصل عليها فرد ما إذا ما تم اختباره باستخدام ذات الاختبار عدداً لا نهائياً من المرات. إذا كانت الدرجات التحصيلية لفرد ما هي X فإن احتمال وقوع درجات هذا الفرد ضمن مجموعة الأفراد الذين تقع درجاتهم الحقيقية على بعد خطأ معياري واحد من X هو واحد إلى ثلاثة.

الدرجات القياسية: وهي مجموعة من الدرجات كدرجات z ودرجات T أو درجات ودرجات الرتب من تسعة (STANINE) تمتلك الوسط الحسابي والإنحراف المعياري المطلوب. يتم حساب الدرجات القياسية عن طريق تحويل الدرجات الأولية إلى درجات z ، ثم ضرب درجات z بالإنحراف المعياري المطلوب، ثم إضافة الوسط الحسابي المطلوب للدرجات المحوَّلة إلى الناتج.

القياس: وهو إجراء اختبار تم بناؤه بعناية فائقة لمجموعة كبيرة من الأفراد تمثل المجتمع تحت ظروف ثابتة وذلك بغرض تحديد معايير للاختبار.

عينة القياس: وهي مجموعة فرعية من المجتمع الذي يستهدفه الاختبار يتم قياس الاختبار على أساسها.

الاختبارات القياسية: وهي اختبارات تم إعدادها بعناية ودقة من قبل محترفين، ثم تم إجراؤها تحت ظروف مثالية وتبعاً لتعليمات ثابتة لمجموعة من الأفراد تمثل المجتمع بغرض التوصل إلى معايير هذه الاختبارات.

الرتب من تسعة (STANINE): مقياس للدرجات المعيارية يتألف من الأعداد ١ إلى ٩ وسطه الحسابي هو ٥ وإنحرافه المعياري هو ٢ تقريباً.

اختيار العينات الطبقي العشوائي: وهو إجراء لأخذ عينات يعتمد على تقسيم المجتمع إلى طبقات (كالنساء والرجال، أو البيض والسود، أو الطبقة الفقيرة والطبقة المتوسطة والطبقة الغنية) ثم إنتقاء عينات عشوائية من كل طبقة مع مراعاة تناسب عدد الأفراد الذين تم اختيارهم من كل طبقة مع حجم الطبقة التي ينتمون إليها.

التقييم التراكمي Summative evaluation: وهو نوع من التقييم يتم في نهاية المقرر أو الوحدة الدراسية ويهدف إلى إعطاء المجموع العام أو القياس النهائي لتحصيل الأفراد.

الاختبار المسأح: وهو اختبار تحصيلي يخالف الاختبار التشخيصي في خصائصه ويهدف إلى تقييم الأداء العام للممتحنين.

درجات T: وهي درجات مشتقة (محوّلة) قياسية مطبّعة وسطها الحسابي ٥٠ وإنحرافها المعياري ١٠، وعلى حين تشابه درجات z درجات T من حيث كون وسطها الحسابي ٥٠ وإنحرافها المعياري ١٠، تبقى درجات z درجات غير مطبّعة.

جدول المواصفات: وهو جدول ذو اتجاهين يتم إعداده كمخطط أو كإطار هيكلية للاختبارات التحصيلية، ويتم فيه تحديد الأهداف السلوكية في الحقول الأفقية وأهداف المحتوى (الموضوع) في الحقول العمودية.

الاختبار المكيف: راجع الاختبار التتابعي.

جداول تايلر - راسل: وهي جداول تقييم لصدق الاختبارات على أساس دلالة المعلومات التي يتم الحصول عليها من الاختبار والتي لا يلعب فيها الحظ أي دور.

تصنيف الأهداف التعليمية: وهي مجموعة من الأهداف المرتبطة بدرس أو اختبار تحصيلي ما مرتبة طبقياً تشمل الطبقات الأعلى فيها كل ما ينتمي إلى الطبقات الأدنى.

تصنيفات السلوك الحركي - النفسي: وهي تصنيفات لأهداف درس أو اختبار ما ضمن المجال الحركي - النفسي.

الاختبار: وهو أية أداة تقييم السلوك أو الأداء لدى الأفراد. وتضم الاختبارات النفسية الاختبارات الإدراكية والعاطفية والحركية - النفسية وغيرها.

القلق المرتبط بالاختبارات Test anxiety: وهو الشعور بالقلق الذي ينتاب الأفراد وقت الاختبار.

التحيز في الاختبارات: يوجد هذا في الاختبارات التي تميز بشكل غير عادل بين مجموعتين أو أكثر من الناس.

ثبات إعادة الاختبار: طريقة لحساب ثبات الاختبار تعتمد على إجراء الاختبار ذاته لنفس المجموعة من الأفراد في وقتين مختلفين ثم حساب الترابط بين درجات الأفراد في المرتين.

البراعة في الاختبارات: وهي عدد من الطرق والإستراتيجيات التي يعتمدها بعض المتحنيين لتحسين درجاتهم في الاختبارات والتي لا ترتبط بمعرفتهم بمادة الاختبار.

الاختبار التجاري: وهو اختبار يقيس معرفة ومهارات المتحن في مجال مرتبط بالوظائف والأعمال المختلفة؛ يُستخدم هذا النوع من الاختبارات في عمليات إنتقاء وتنسيب الموظفين ومنح التراخيص المهنية.

المتخلفين عقلياً الممكن تدريبهم (TMR): وهم الأطفال الذين تعتبر درجة تخلفهم العقلي (تتراوح معدلات ذكاء هؤلاء الأطفال بين ٢٦ و ٥٠) درجة لا يمكن معها تعلم القراءة والكتابة إلا أنها تسمح بأداء المهام التي لا تحتاج إلى مهارة تحت المراقبة.

فقرات الصواب والخطأ: وهي نوع من الفقرات الاختبارية تحتوي على بيانات قد تكون صحيحة أو خاطئة.

الدرجات الفعلية (الحقيقية): معدل افتراضي لدرجات فرد ما يعبر عن معرفة الفرد الفعلية لمادة الاختبار. تنص نظرية الاختبارات على أن حساب الدرجات الفعلية للفرد يتم عن طريق حساب الوسط الحسابي لتوزيع درجات هذا الفرد إذا ما تم اختياره باستخدام الاختبار ذاته عدداً لا نهائياً من المرات.

الصدق: وهو درجة قياس أداة قياس ما للمتغير الذي تم تصميمها لقياسه. ومن أنواع صدق الاختبارات: صدق المحتوى (ويتم حسابه عن طريق تحليل محتوى الاختبار)، والصدق التنبؤي والتلازمي (ويتم حسابه عن طريق الربط بين درجات

الاختبار ودرجات مقياس المحك)، وصدق المفهوم (ويتم حسابه عن طريق دراسة مستفيضة لمدى قياس اختبار ما لمفهوم نفسي ما).

معامل الصدق: وهو الارتباط بين درجات المتغير التنبؤي ودرجات متغير المحك.

تعميم الصدق: وهو تطبيق أدلة الصدق على مواقف غير تلك التي تم منها الحصول على هذه الأدلة.

التغاير (التنوع): وهو توزع أو انحراف عدد من الدرجات حول قيمها الوسطية.

التباين: وهو مقياس لتغاير (تنوع) درجات اختبار ما يتم حسابه كمجموع مربعات إنحرافات الدرجات الخام من الوسط الحسابي مقسماً على عدد ينقص عن مجموع عدد معدلات الدرجات بمقدار ١؛ كذلك مربع الانحراف المعياري.

القدرة اللغوية: القدرة على فهم الكلمات والعلاقات بينها في لغة ما؛ كذلك القدرة على حل المسائل المرتبطة بالكلمات.

الاختبار اللفظي (اللفظي): وهو اختبار تعليماته لغوية وتتطلب الإجابة الشفهية أو الكتابية على أسئلته كلمات أو أعداد.

قائمة الرصد للجماعات: وهي قائمة رصد يستخدمها المدرس لتقييم أداء مجموعة من الطلاب في آن واحد.

اختبار عينة العمل: وهو اختبار تشابه فقراته المهام التي سيقوم بها الممتحن في حال نجاحه في الحصول على الوظيفة التي تقدم إليها بطلب توظيف.

درجات z: أي من مجموعة من الدرجات المشتقة تتراوح بين $-\infty$ و $+\infty$ ويتم حسابها من المعادلة: $z = (\text{الدرجات الخام} - \text{الوسط الحسابي}) / \text{الانحراف المعياري}$ ، وذلك لكل معدلات الدرجات الخام. وفي التوزيع الطبيعي للدرجات، تقع ٩٩٪ من الحالات بين $z = 2-$ و $z = 2+$.

منطقة التطور الممكن: وهي المسافة أو الفرق بين المستوى التطوري الفعلي للطفل (وهي مستوى التطور الذي حققه الطفل والذي تم تقديره عن طريق اختبار قياسي) وإمكانية تطوره (وهي درجة الكفاءة التي يمكن أن يصل إليها الطفل مع المساعدة).

obeikandi.com

المراجع

- Aiken, L. R. (1970). Scoring for partial knowledge on the generalized rearrangement test item. *Educational and Psychological Measurement*, 30, 87-94.
- Aiken, L. R. (1983). Determining grade boundaries on classroom tests. *Educational and Psychological Measurement*, 43, 759-762.
- Aiken, L. R. (1987). Testing with multiple-choice items. *Journal for Research and Development in Education*, 20(4), 44-58.
- Aiken, L. R. (1988). A program for computing the reliability and maximum reliability of a weighted composite. *Educational and Psychological Measurement*, 48, 703-706.
- Aiken, L. R. (1989a). Some simple item analysis statistics and significance tests. *Journal for Research and Development in Education*, 22(2), 42-48.
- Aiken, L. R. (1989b). Analyzing optional test items. *Educational and Psychological Measurement*, 49, 783-787.
- Aiken, L. R. (1991). Detecting, understanding, and controlling for cheating on tests. *Research in Higher Education*, 32, 725-736.
- Aiken, L. R. (1996). *Rating scales and checklists: Evaluating behavior, personality and attitudes*. New York: Wiley.
- Aiken, L. R. (1997a). *Psychological testing and assessment* (9th ed.). Needham Heights, MA: Allyn & Bacon.
- Aiken, L. R. (1997b). *Questionnaires and inventories: Surveying opinions and assessing personality*. New York: Wiley.
- Airasian, P. W., & Terrasi, S. (1994). Test administration. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 11, pp. 6311-6315). Tarrytown, NY: Elsevier.
- Allison, D. E. (1984). The effect of item-difficulty sequence, intelligence, and sex on test performance, reliability, and item difficulty and discrimination. *Measurement and Evaluation in Guidance*, 16, 211-217.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, 47, 1597-1611.

- American Psychological Association, Committee on Professional Standards, & Committee on Psychological Tests and Assessment. (1986). *Guidelines for computer-based tests and interpretations*. Washington, DC: American Psychological Association.
- Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Macmillan.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). New York: Macmillan.
- Angoff, W. H., & Ford, S. F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement*, 10, 95-106.
- Bender, L. (1938). *A visual motor gestalt test and its clinical use* (Research Monograph No. 3). New York: American Orthopsychiatric Association.
- Binet, A., & Simon, T. (1905). Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *L'Année Psychologique*, 11, 191-244.
- Black, H. (1962). *They shall not pass*. New York: Morrow.
- Blakley, B. R., Quinones, M. A., Crawford, M. S., & Jago, I. A. (1994). The validity of isometric strength tests. *Personnel Psychology*, 47, 247-274.
- Bloom, B. S., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I. The cognitive domain*. New York: McKay.
- Bloom, B. S., & Krathwohl, D. R. (1984). *Taxonomy of educational objectives: Handbook I. Cognitive domain*. White Plains, NY: Longman.
- Bredemeier, M. (1991, November/December). IQ test ban for blacks called unconstitutional. *California Association of School Psychologists Today*, 22-23.
- Bridgman, P. W. (1927). *The logic of modern physics*. New York: Macmillan.
- Brigham, C. C. (1923). *A study of American intelligence*. Princeton, NJ: Princeton University Press.
- Brigham, C. C. (1930). Intelligence tests of immigrant groups. *Psychological Review*, 37, 158-165.
- Budoff, M. (1987). The validity of learning potential assessment. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 52-81). New York: Guilford Press.
- Camilli, G. (1979). *A critique of the chi square method for assessing item bias*. Unpublished paper, Laboratory of Educational Research, University of Colorado, Boulder.
- Campione, J. C., & Brown, A. L. (1987). Learning ability and transfer propensity as sources of individual differences in intelligence. In P. H. Brooks, R. D. Sperber, & C. McCauley (Eds.), *Learning and cognition in the mentally retarded* (pp. 265-294). Baltimore: University Park Press.
- Carey, L. M. (1994). *Measuring and evaluating school learning* (2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Carroll, J. B. (1973). The aptitude-achievement distinction: The case of foreign language aptitude and proficiency. In D. R. Green (Ed.), *The aptitude-achievement distinction*. Monterey, CA: CTB/McGraw-Hill.
- Carroll, J. B. (1978). On the theory-practice interface in the measurement of intellectual abilities. In P. Suppes (Ed.), *Impact of research on education* (pp. 1-105). Washington, DC: National Academy of Education.
- Cattell, J. M. (1890). Mental tests and measurements. *Mind*, 15, 373-381.

- Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple sensory systems. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory*. New York: Guilford Press.
- Collet, L. S. (1971). Elimination scoring: An empirical evaluation. *Journal of Educational Measurement*, 8, 209-214.
- Conoley, J. C., & Impara, J. C. (Eds.). (1995). *The twelfth mental measurements yearbook*. Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.
- Conoley, J. C., & Kramer, J. J. (Eds.). (1989). *The tenth mental measurements yearbook*. Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.
- Cooley, W. W., & Lohnes, P. R. (1968). *Predicting development of young adults* (Project TALENT Five-year Follow-up Studies, Interim Report 5, American Institutes for Research). Washington, DC: U.S. Office of Education.
- Crowl, T. K., & McGinitie, W. H. (1974). The influence of students' speech characteristics on teachers' evaluations of oral answers. *Journal of Educational Psychology*, 66, 304-308.
- Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Needham Heights, MA: Allyn & Bacon.
- David, P. (1981, November 27). Multiple-choice tests under fire: United States. *Times Education Supplement*, 3413(11).
- Davis, C. J. (1980). *Perkins-Binet tests of intelligence for the blind*. Watertown, MA: Perkins School for the Blind.
- Delahunty, R. J. (1988). Perspectives on within-group scoring. *Journal of Vocational Behavior*, 33, 463-477.
- Ebel, R. L. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Edouard, L. J., & Harris, F. T. C. (1983). Multiple-choice questions: Fact, fallacy and fancy. *International Journal of Mathematical Education in Science and Technology*, 14, 685-691.
- Egeland, B. (1985). Review of Wisconsin card sorting test. In J. V. Mitchell (Ed.), *The ninth mental measurements yearbook* (pp. 1746-1747). Lincoln: University of Nebraska Press.
- Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika*, 49, 175-186.
- Evans, W. (1984). Test wiseness: An examination of cue-using strategies. *Journal of Experimental Education*, 52, 141-144.
- Eyde, L., Robertson, G. J., Krug, S., & Others. (1993). *Responsible test use: Case studies for assessing human behavior*. Washington, DC: American Psychological Association.
- Fabiano, E. (1989). *Index to tests used in educational dissertations*. Phoenix, AZ: Oryx Press.
- Feuerstein, R., Rand, Y., Jensen, M. R., Kaniel, S., & Tzurriel, D. (1987). Prerequisites for assessment of learning potential: The LPAD model. In C. S. Lidz (Ed.), *Dynamic assessment: An interactive approach to evaluating learning potential* (pp. 35-51). New York: Guilford Press.
- Fremer, J., Diamond, E. E., & Camara, W. J. (1989). Developing a code of fair testing practices in education. *American Psychologists*, 44, 1062-1067.

- Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.
- Gerlach, V. S., & Sullivan, H. J. (1967). *Constructing statements of outcomes*. Inglewood, CA: Southwest Laboratory for Educational Research and Development.
- Ghiselli, E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461-477.
- Goldman, B. A., & Mitchell, D. F. (Eds.). (1990). *Directory of unpublished experimental mental measures* (Vol. 5). Dubuque, IA: Wm. C. Brown.
- Goldman, B. A., & Osborne, W. L. (Eds.). (1985). *Directory of unpublished mental measures* (Vol. 4). New York: Human Sciences Press.
- Green, B. F. (1992). Testing technology. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 4, pp. 1410-1414). New York: Macmillan.
- Green, D. R. (1992). Test construction. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 4, pp. 1400-1406). New York: Macmillan.
- Green, J. A. (1975). *Teacher-made tests* (2nd ed.). New York: Harper & Row.
- Green, K. (1984). Effects of item characteristics on multiple-choice item difficulty. *Educational and Psychological Measurement*, 44, 551-561.
- Guerra, E. L., Abramson, D. A., & Newmark, M. (1964). New York City foreign language oral ability rating scale. *Modern Language Journal*, 48, 486-489.
- Hakstian, A. R., & Kansup, W. (1975). Comparison of several methods of assessing partial knowledge in multiple-choice tests: Testing procedures. *Journal of Educational Measurement*, 12, 231-239.
- Hammill, D. D., Brown, L., & Bryant, B. R. (1992). *A consumer's guide to tests in print* (2nd ed.). Austin, TX: Pro-Ed.
- Harrington, R. G. (Ed.). (1986). *Testing adolescents*. Kansas City, MO: Test Corporation of America.
- Harrow, A. J. (1972). *A taxonomy of the psychomotor domain: A guide for developing behavioral objectives*. New York: McKay.
- Hartigan, J. A., & Wigdor, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press.
- Herbert, E. A. (1992). Portfolios invite reflection—from both students and staff. *Educational Leadership*, 49(8), 58-61.
- Herman, J. L. (1994). Item writing techniques. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 5, pp. 3061-3066). Tarrytown, NY: Elsevier.
- Hoffman, B. (1962). *The tyranny of testing*. New York: Crowell-Collier.
- Hogan, J., & Quigley, A. (1994). Effects of preparing for physical ability tests. *Public Personnel Management*, 23(1), 85-104.
- Hopkins, K. D., & Bracht, G. H. (1975). Ten-year stability of verbal and nonverbal IQ scores. *American Educational Research Journal*, 12, 469-477.
- Hughes, S. (1995). Review of Denver: II. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 263-265). Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.
- Hunter, J. E. (1983). *The dimensionality of the General Aptitude Tests Battery (GATB) and the dominance of general factors over specific factors in the prediction of job performance* (USES Test Research Report No. 44). Washington, DC: U.S. Department of Labor.

- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.
- Intelligence and its measurement: A symposium. (1921). *Journal of Educational Psychology*, 12, 123-147, 195-216.
- Jensen, A. R. (1982a). Reaction time and psychometric *g*. In H. J. Eysenck (Ed.), *A model for intelligence*. Berlin: Springer-Verlag.
- Jensen, A. R. (1982b). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in research in intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Kaufman, A. S. (1978). Review of Columbia Mental Maturity Scale. In O. K. Buros (Ed.), *The eighth mental measurements yearbook* (Vol. 1, pp. 299-301). Highland Park, NJ: Gryphon Press.
- Kaufman, A. S., & Kaufman, N. L. (1983a). *Kaufman Assessment Battery for Children: Administration and scoring manual*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S., & Kaufman, N. L. (1983b). *Kaufman Assessment Battery for Children: Interpretive manual*. Circle Pines, MN: American Guidance Service.
- Keyser, D. J., & Sweetland, R. C. (Eds.). (1984-1994). *Test critiques* (Vols. 1-10). Austin, TX: Pro-Ed.
- Kiess, H. O. (1996). *Statistical concepts for the behavioral sciences* (2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Klimko, I. P. (1984). Item arrangement, cognitive entry characteristics, sex, and test anxiety as predictors of achievement examination performance. *Journal of Experimental Education*, 52, 214-219.
- Knobloch, H., & Pasamanick, B. (Eds.). (1974). *Gesell and Amatruda's developmental diagnosis* (3rd ed.). New York: Harper & Row.
- Knobloch, H., Stevens, F., & Malone, A. (1987). *Manual of developmental diagnosis: The administration and interpretation of the Revised Gesell and Amatruda Developmental and Neurologic Examination*. Houston, TX: Developmental Evaluation Materials.
- Kramer, J. J., & Conoley, J. (1992). *The eleventh mental measurements yearbook*. Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.
- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: Handbook II. The affective domain*. New York: McKay.
- Krug, S. E. (1993). *Psychware sourcebook* (4th ed.). Champaign, IL: MetriTech.
- Kubiszyn, T., & Borich, G. (1990). *Educational testing and measurement*. Glenview, IL: Scott, Foresman.
- Lacks, P. (1984). *Bender-Gestalt screening for brain dysfunction*. New York: Wiley.
- Larson, G. E. (1994). Armed Services Vocational Aptitude Battery. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 121-124). New York: Macmillan.
- Levine, M. (1976). The academic achievement test: Its historical context and social functions. *American Psychologist*, 31, 228-238.
- Levy, P., & Goldstein, H. (1984). *Tests in education: A book of critical reviews*. New York: Academic Press.
- Linn, R. L. (1992). Achievement testing. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1-12). New York: Macmillan.

- Little, E. B. (1962). Overcorrection for guessing in multiple-choice test scoring. *Journal of Educational Research*, 55, 245-252.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719-748.
- Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore: William & Wilkins.
- McClelland, D. C. (1961). *The achieving society*. Princeton, NJ: Van Nostrand.
- McDaniel, E. (1994). *Understanding educational measurement*. Madison, WI: Brown & Benchmark.
- McNemar, Q. (1964). Lost: Our intelligence? Why? *American Psychologist*, 19, 871-882.
- Mislevy, R. J., & Stocking, M. L. (1989). A consumer's guide to LOGIST and BILOG. *Applied Psychological Measurement*, 13, 57-75.
- Mountain, M., & Snow, W. (1993). Wisconsin Card Sorting Test as a measure of frontal pathology: A review. *Clinical Neuropsychologist*, 7, 108-118.
- Murphy, K. R. (1984). Review of Armed Services Vocational Aptitude Battery. In D. Keyser & R. Sweetland (Eds.), *Test critiques* (Vol. 1, pp. 61-69). Kansas City, MO: Test Corporation of America.
- Murphy, K. R., & Davidshofer, C. O. (1994). *Psychological testing: Principles and applications* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Murphy, L. L., Conoley, J. C., & Impara, J. C. (Eds.). (1994). *Tests in print: IV*. Lincoln: University of Nebraska and Buros Institute of Mental Measurements.
- Nixon, J. E., & Jewett, A. E. (1980). *An introduction to physical education* (9th ed.). Philadelphia: Saunders.
- Norušis, M. J. (1992a). *SPSS/PC+ base system user's guide version 5.0*. Chicago, IL: SPSS.
- Norušis, M. J. (1992b). *SPSS/PC+ advanced statistics version 5.0*. Chicago, IL: SPSS.
- Norušis, M. J. (1992c). *SPSS/PC+ professional statistics version 5.0*. Chicago, IL: SPSS.
- Pagano, R. R. (1994). *Understanding statistics* (4th ed.). Minneapolis/St. Paul: West.
- Parents in Action on Special Education (PASE) v. Joseph P. Hannon, No. 74C 3586 (N.D. III. 1980).
- Paterson, D. G., Elliott, R. M., Anderson, L. D., Toombs, H. A., & Heidbreder, E. (1930). *Minnesota Mechanical Ability Tests*. Minneapolis: University of Minnesota Press.
- Paulson, F. L., & Paulson, P. R. (1992). *The making of a portfolio*. Unpublished manuscript. (Available from authors, 6800 Gable Parkway, Portland, OR 97225)
- Pellegrino, J. (1992). Intelligence and aptitude measurement. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 2, pp. 627-637). New York: Macmillan.
- Pellegrino, J. (1994). Abilities and aptitudes. In T. Husén & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed.) (Vol. 1, pp. 1-6). Tarrytown, NY: Elsevier.
- Plake, B. S., Anson, C. J., Parker, C. S., & Lowry, S. R. (1982). Effects of item arrangement, knowledge of arrangement, test anxiety and sex on test performance. *Journal of Educational Measurement*, 19, 49-57.

- Platt, J. R. (1961). On maximizing the information obtained from science examinations. *American Journal of Physics*, 29, 111-122.
- Preston, R. C. (1964). Ability of students to identify correct responses before reading. *Journal of Educational Research*, 58, 181-183.
- Pyle, W. H. (1913). *The examination of school children*. New York: Macmillan.
- Raymond, M. R., Webb, L. C., & Houston, W. M. (1991). Correcting performance-rating errors in oral examinations. *Evaluation and the Health Professions*, 14(1), 100-122.
- Reckase, M. D., Ackerman, T. A., & Carlson, J. E. (1988). Building a unidimensional test using multidimensional items. *Journal of Educational Measurement*, 25, 193-203.
- Reid, F. (1977). Alternative scoring formula for multiple-choice and true-false tests. *Journal of Educational Research*, 70, 335-339.
- Reitan, R. M. (1966). A research program on the psychological effects of brain lesions in human beings. In N. R. Ellis (Ed.), *International research in mental retardation* (Vol. 1, pp. 153-218). New York: Academic Press.
- Resnick, D. (1982). History of educational testing. In A. K. Wigdor & W. R. Garner (Eds.), *Ability testing: Uses, consequences, and controversies, Part 2* (pp. 173-194). Washington, DC: National Academic Press.
- Ross, C. C., & Stanley, J. C. (1954). *Measurement in today's schools* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Rowley, G. L. (1974). Which examinees are most favoured by the use of multiple-choice tests? *Journal of Educational Measurement*, 11, 15-23.
- Safrit, M. J., & Wood, T. M. (1995). *Introduction to measurement in physical educational and exercise science* (3rd ed.). St. Louis: Mosby.
- Sample, B. M. (1992). *Performance assessment: An international experiment* (Report No 22-Caep-06). Princeton, NJ: Educational Testing Service.
- Sattler, J. M. (1988). *Assessment of children* (3rd ed.). San Diego: Sattler.
- Savitz, F. R. (1985). Effects of easy examination questions placed at the beginning of science multi-choice examinations. *Journal of Instructional Psychology*, 12, 6-10.
- Scheuneman, J. (1979). A new method for assessing bias in test items. *Journal of Educational Measurement*, 16, 143-152.
- Schmidt, F. L. (1985). Review of Wonderlic Personnel Test. In J. V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (Vol. 2, pp. 1755-1757). Lincoln: Buros Institute of Mental Measurements, University of Nebraska.
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62, 529-540.
- Schmidt, F. L., & Hunter, J. E. (1978). Moderator research and the law of small numbers. *Personnel Psychology*, 31, 215-232.
- Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. *American Psychologist*, 36, 1128-1137.
- Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1982). Progress in validity generalization: Comments on Callender and Osbury and further developments. *Journal of Applied Psychology*, 67, 835-845.
- Schmidt, F. L., Hunter, J. E., Pearlman, K., Hirsch, H. R., Sackett, P. R., Schmitt, N., Tenopyr, M. L., Kehoe, I., & Zedeck, S. (1985). Forty questions about validity

- generalizations and meta-analysis with commentaries. *Personnel Psychology*, 37, 407-422.
- Schoenfeldt, L. F. (1985). Review of Wonderlic Personnel Test. In J. V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (Vol. 2, pp. 1757-1758). Lincoln: Buros Institute of Mental Measurements, University of Nebraska.
- Shaha, S. H. (1984). Matching tests: Reduced anxiety and increased test effectiveness. *Educational and Psychological Measurement*, 44, 869-881.
- Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences* (3rd ed.). Needham Heights, MA: Allyn & Bacon.
- Shea, C. (1994, September 7). "Gender gap" on examinations shrank again this year. *The Chronicle of Higher Education*, p. A54.
- Simpson, E. J. (1966). The classification of educational objectives, psychomotor domain. *Illinois Teacher of Home Economics*, 10, 110-114.
- Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist*, 27(1), 5-32.
- Sperry, R. W. (1968). Hemisphere disconnection and unity in conscious awareness. *American Psychologist*, 23, 723-733.
- Stalnaker, J. L. (1951). The essay type of examination. In E. Lindquist (Ed.), *Educational measurement* (pp. 495-530). Washington, DC: American Council on Education.
- Starch, D., & Elliott, E. C. (1913). Reliability of grading work in mathematics. *School Review*, 21, 254-259.
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. Columbus, OH: Macmillan.
- Stoloff, M. L., & Couch, J. V. (Eds.). (1992). *Computer use in psychology: A directory of software* (3rd ed.). Washington, DC: American Psychological Association.
- Strand, B. N., & Wilson, R. (1993). *Assessing sport skills*. Champaign, IL: Human Kinetics.
- Sweetland, R. C., & Keyser, D. J. (Eds.). (1991). *Tests* (3rd ed.). Austin, TX: Pro-Ed.
- Swiercinsky, D. P. (Ed.). (1985). *Testing adults*. Kansas City, MO: Test Corporation of America.
- Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, 51, 567-577.
- Thorndike, R. L. (1985). The central role of general ability in prediction. *Multivariate Behavioral Research*, 20, 241-254.
- Thorndike, R. L., & Hagen, E. P. (1959). *Ten thousand careers*. New York: Wiley.
- U.S. Department of Defense. (1993). *ASVAB 18/19 technical manual*. North Chicago, IL: U.S. Military Entrance Processing Command.
- U.S. Department of Defense. (1994). *Technical manual for the ASVAB 18/19 career exploration programs*. North Chicago, IL: U.S. Military Entrance Processing Command.
- Vale, C. D. (1985). ASCAL: Item parameter estimation program [Computer program]. St. Paul, MN: Assessment Systems.
- Vermont State Board of Education. (1991). *Looking beyond "the answer": Vermont's mathematics portfolio assessment program, pilot year, 1990-1991*. Montpelier: Vermont State Board of Education.
- Vernon, P. E. (1960). *The structure of human abilities* (Rev. ed.). London: Methuen.

- Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: MIT Press. (Original work published 1934)
- Vygotsky, L. S. (1978). In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds. & Trans.), *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Weaver, S. J. (Ed.). (1984). *Testing children*. Kansas City, MO: Test Corporation of America.
- Welsh, J. R., Kucinkas, S. K., & Curran, L. T. (1990). *Armed Services Vocational Aptitude Battery (ASVAB): Integrative review of validity studies*. San Antonio, TX: Brooks Air Force Base.
- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.
- Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review Monograph Supplement*, 3(No. 6).
- Wolf, D. P. (1989). Portfolio assessment: Sampling student work. *Educational Leadership*, 46(7), 35-36.
- Woo, E. (1997, May 7). Few signs found of academic gender gap. *Los Angeles Times*, pp. A1, A33.
- Wright, B. D., & Linacre, M. (1991). *A user's guide to BIGSTEPS*. Chicago: MESA Press.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: MESA Press.
- Yerkes, R. M. (1919). Report of the Psychology Committee of the National Research Council. *Psychological Review*, 26, 83-149.
- Yerkes, R. M. (Ed.). (1921). Psychological examining in the United States army. *Memoirs of the National Academy of Sciences*, 15.
- Zeidner, J., & Johnson, C. (1992). *Is personnel classification a concept whose time has passed?* Presented at the ARI Conference on Selection and Classification in the Military, Alexandria, VA.