٣

تصميم ووضع وتصحيح الاختبارات

تتوازى المعلومات عن تصميم الاختبارات الموجودة في القسم الأول من هذا الفصل مع معلومات الفصل الثاني من كتاب سلالم التقدير وقوائم الرصد (أيكين، ١٩٩٦) الذي يعالج قضية تصميم ووضع سلالم التقدير وقوائم الرصد. وكذلك مع معلومات الفصل الثاني من كتاب الاستبانات ومقاييس السمات (أيكين، ١٩٩٧). ومن باب التذكير، إن الكتابين المذكورين آنفاً هما الجزءان الأول والثاني من هذه السلسلة. يناقش هذا الفصل أيضاً ثلاثة عناصر أخرى من عناصر الاختبار هي: إجراء الاختبار وتصحيحه ثم تفسير نتائجه. وتساعد برامج الفئة ب الموجودة على القرص المرافق لهذا الكتاب على فهم وإنجاز هذه المهام، إذ يساعد برنامج B-1 على تصميم اختبارات موضوعية، ويساعد برنامج B-2 على إجراء ووضع درجات الاختبارات الموضوعية، بينما يساعد برنامج B-3 على عملية "التصحيح في حالة التخمين" Guessig في الاختبارات الموضوعية. أما برنامج B-4 فيساعد على وضع درجات أسئلة المطابقة Matching أو إعادة الترتيب. ويستخدم برنامج B-5 في وضع درجات الأسئلة التي تتطلب ترتيب

obeikandl.com

الفقرات. أما برنامج B-6 فيستخدم في وضع الدرجات وفق طريقة كاجوري المعدلة (Cajori)، بينما يفيد برنامج B-7 والذي يعطي تقديراً لاحتمالية إرتكاب ممتحنين جالسين على مقربة من بعض نفس الأخطاء في التعرف على حالات الغش في الاختبارات.

تصميم الاختبارات

لا تتطلب الغالبية الساحقة من الاختبارات النفسية والتعليمية تصميماً شديد الدقة أو إعداداً شاقاً وصعباً إذ أن معظم هذه الاختبارات هي اختبارات وليدة اللحظة يضعها المدرسون وغيرهم من المقيمين الذين هم بحاجة إلى وسيلة مختصرة وبسيطة تمكنهم من تقييم تحصيل الفرد العلمي أو مكانته. وغالباً ما تستأثر عملية التدريس بجل اهتمام المدرسين فلا يعيرون اهتماماً مناسباً لعملية قياس ما إذا كانت الأهداف التعليمية قد تم تحقيقها أوإلى أي درجة تم تحقيقها. وعليه، لا يقضي هؤلاء المدرسون عادةً وقتاً طويلاً في وضع اختبارات مقالية أو اختبارات موضوعية تقيس درجة تحصيل الطلاب. وعلى عكس الاختبارات التي يضعها المدرسون والتي تعد اختبارات "الفرصة الواحدة" للطالب، تعد الاختبارات وأدوات التقييم التي يضعها محترفو تصميم الاختبارات والتي تسوقها شركات تجارية للاختيارات.

يتوفر عدد كبير من الاختبارات التي تقيم المهارات الإدراكية والعاطفية والنفسية – الحركية، كما يتم تصميم عدد كبير من الاختبارات كل يوم والتي قد تنشر أو لا تنشر. وتختلف طرق تصميم الاختبارات إلى حد ما في نوع الاختبار ولكن في الإحالة.

يجب أن يبدأ الاختبار ببيان الفرض من الاختبار، ويتضمن هذا البيان بشكل عام إجابات للأسئلة التالية:

١- ما هي المتغيرات التي يهدف الاختبار قيد التصميم إلى قياسها؟

٢- من سيتم اختباره؟

٣- ما هي الظروف التي سيجرى ضمنها الاختبار؟

٤- كيف سيتم وضع درجات الاختبار؟

٥- ما هي الجهة التي سوف يتم تبليغها بنتائج الاختبار وكيف سيتم استخدام هذه النتائج؟

٦- ما هي الموارد المالية المتوفرة لدعم عمليات تصميم وإجراء الاختبارات ثم وضع درجاتها وتحليل نتائجها؟

وللإجابة على السؤال الأول، يتوجب على مصمم الاختبار أن يحدد ما إذا كان المطلوب هو قياس التحصيل أو الذكاء أو الاستعدادات الخاصة، إذ غالباً ما تركز الاختبارات التحصيلية على ما قد تعلمه الممتحن ضمن مجال مهاري أو معرفي معين في حين تتطلع اختبارات الاستعداد إلى المستقبل أو ما يمكن لفردٍ ما أن يتعلمه إذا توفرت له الظروف الموائمة. ولهذا السبب تتناول معظم فقرات الاختبارات التحصيلية مادةً علمية محددة تم تدريسها بشكل تقليدي ولا تتطرق إلى التجارب الإنسانية العامة. ويجب عند تصميم الاختبارات التحصيلية التي تختبر مواد علمية محددة كالقراءة والرياضيات واللغة الانجليزية أو اللغات الأجنبية إعداد قائمة بمواصفات المادة العلمية موضع الاختبار، ثم يجب إتخاذ قرارات تتعلق بأنواع الفقرات الاختبارية التي يمكنها أن تمثل المادة العلمية بشكل جيد وتقيسها بشكل فعال.

وتفيد الإجابة على السؤالين ٢ و٣ المتعلقين بمن سوف يتم اختباره وما هي الظروف التي سيجري ضمنها الاختبار في تحديد نوعية الفقرات الاختبارية المستخدمة (فقرات موضوعية أو مقالية أو أدائية) في الاختبار وتحديد طول

الاختبار وشكل تقديم الاختبار (على شكل كتيب أو تجربة عملية أو مقاطع شفهية) واللغة المستخدمة في الفقرات المطبوعة أو المحكية. وغالباً ما تكون الاختبارات الفردية مؤلفة من عدد من وسائل الاتصال (شفهية وكتابية وتطبيقية) التي تتطلب تعليمات شفهية مختلفة لكل قسم أو مجموعة من الفقرات. أما الاختبارات الجماعية فيتم تقديمها على شكل كتيب وترافق التعليمات المكتوبة في الكتيب تعليمات شفهية بسيطة من القائمين على الاختبار. ترتبط القرارات المتخذة بشأن محتوى الاختبار وشكله على الوقت المتوفر لإجراء الاختبار، إذ يمكن خلال ثلاث ساعات طرح عدد أكبر من الأسئلة واختبار عدد أكبر من المتغيرات مما يمكن طرحه خلال ساعتين إلا أن المشكلة تكمن في كون الاختبارات التي تستمر لأكثر من ساعتين اختبارات تؤثر سلبياً على تركيز الممتحن ورغبته في الإجابة على الأسئلة الأمر الذي يحول دون قياس قدرة الممتحن الحقيقة بشكل دقيق. ويتناول هذا الفصل من الكتاب الطرق المحوسبة في إجراء الاختبارات التي يجب أن تؤخذ بعين الاعتبار عند تصميم الاختبارات.

وتلعب طريقة وضع درجات الاختبار (باليد أو عن طريق الماسحة الالكترونية البصرية أو عن طريق إعطاء الممتحن فرصة وضع درجاته بنفسه) دوراً هاماً في عملية تصميم الاختبارات. ففي معظم الاختبارات الفردية يكون القائم على الاختبار وواضع الدرجات (أو المصحح) هما الشخص ذاته وغالباً ما تتم عملية التصحيح أثناء عملية الاختبار، أما في حالة الاختبارات الجماعية فلا يتم وضع الدرجات إلا بعض الانتهاء من الاختبار بكامله حيث يقوم القائم على الاختبار أو موظف مسؤول أو أي شخص كفؤ بوضع الدرجات. وتعتمد درجة تفصيل المعلومات المكتوبة في كشف الدرجات على الجهة المعنية بقراءة هذه المعلومات وعلى طريقة الحصول عليها. إذ يجب أن تتم كتابة المعلومات بطريقة مفهومة

لكل من لديه الحق بالإطلاع على النتائج ولديه القدرة على استخدام المعلومات الناتجة لما فيه مصلحة الممتحن.

ويمكن طرح عدد من الأسئلة المرتبطة بالموارد المالية المتوفرة لتصميم الاختبار: هل سيتقاضى مصمم الاختبار أجراً وما هو مقدار هذا الأجر؟ هل من الأفضل والأوفر تقديم الاختبار في قالب معين أو قالب آخر؟ ما هي الظروف المثلى اقتصادياً لإجراء الاختبار وتصحيحه؟

التخطيط وتحديد مواصفات الاختبارات التحصيلية

تختلف خطة ومواصفات الاختبارات التحصيلية عن خطط ومواصفات اختبارات القدرات الإدراكية العامة أوالخاصة. وتختلف مواصفات اختبار التحصيل باختلاف نوع الاختبار من حيث كونه اختبار خلاصي (أو اختبار نهائي يأتي في نهاية الوحدة الدراسية أو المقرر بأكمله) أو اختبار تكويني Formative (أي اختبار يجرى دورياً يتم على أساسه تعديل المقرر أو طريقة التدريس). وعلى الرغم من أن مواصفات الاختبار تختلف باختلاف موضوع الاختبار، يبقى إعداد جدول بمواصفات مفصلة لمحتوى الاختبار أمراً لا بد منه لضمان تحقيق الاختبار لأهدافه.

درج الباحثون في مجال التعليم خلال السبعينيات والستينيات على وضع أهداف محددة بدقة جداً للمقررات التعليمية وقد تم استخدام دلائل التصنيف مثل التي وضعها بلوم وكارثول (١٩٥٦) أو غرلاش وسوليفان (١٩٦٧) أو إيبيل (١٩٧٩) في تصميم أدوات التعليم والتقييم. ويتألف دليل بلوم وكارثول (١٩٥٦، ١٩٨٤)، وهو الدليل الأكثر شعبية للأهداف التعليمية، من ترتيب هرمي ينقسم إلى ٦ مجموعات هي: المعرفة والفهم والتطبيق والتحليل والتركيب والتقييم. لا تعد هذه المجموعات الستة التي تم ترتيبها من الأبسط إلى الأعقد ومجموعات

غير منتمية ولكنها شاملة بشكل متبادل بحيث أن الأهداف التعليمية المنتمية إلى مجموعات في الطبقات الدنيا تم تضمينها في مجموعات الطبقات العليا بالضرورة.

تم استخدام دلائل التصنيف مثل تلك التي وضعها بلوم وكارثول (١٩٥٦) لتحديد أحد أبعاد جداول مواصفات الاختبارات. وتُصنف أهداف المقررات أو الوحدات التدريسية غالباً على أنها جداول ذات بعدين، يتألف البعد الأول (الحقول الأفقية) منها من مجموعات الأهداف السلوكية والبعد الثاني من مجموعات أهداف محتوى الاختبار (الحقول العمودية). فعلى سبيل المثال، يمكن لجدول مواصفات اختباريقييم عمليتي إعداد وإجراء الاختبارات التحصيلية أن يتألف من أربعة حقول أفقية تحدد الأهداف السلوكية للاختبار (معرفة المصطلحات، معرفة الحقائق المحددة، الاستيعاب، التطبيق) وخمسة أعمدة لمواصفات المحتوى (الإعداد للاختبار، وضع الاختبار، إجراء الاختبار، وضع الدرجات، تحليل فقرات الاختبار). وعند نقطة تقاطع الحقول الأفقية والعمودية، تتوفر قوائم تحتوي وصفاً مختصراً لفقرات الاختبار التي تتناول موضوعاً ما وذكراً لعدد هذه الفقرات.

وعلى الرغم من أن معظم المدرسين وغيرهم من واضعي الاختبارات من غير المحترفين لا يضعون جدولاً مفصلاً بمواصفات الاختبارات الخاصة بظرف ما التي يستخدمونها، يبقى من الحكمة أن يكون لدى واضع الاختبار خطة دقيقة للاختبار وغالباً ما تحتوي الخطة التفصيلية لمحتوى الاختبارعلى قائمة بالمجموعات الأساسية والفرعية لكل موضوع يجب أن تتم تغطيته وعدد (أو نسبة) الفقرات الاختبارية التي تم تخصيصها لكل من هذه المواضيع. وبهذه الطريقة يضمن واضع الاختبار تغطية الاختبار لكافة المواضيع المراد اختبار تحصيل الطلاب فيها كما يضمن واضع الاختبار أن الاختبار مناسب من حيث الطول والشكل للممتحنين وللوقت والموارد المتاحة. ومن الحكمة إعداد فقرات

اختبار تزيد ٢٠٪ عن العدد المطلوب للاختبار في شكله النهائي، ويتم الاختيار النهائي للفقرات بعد اتمام عمليات المراجعة وبعد الإطلاع على نتائج تجربة الفقرات الاختبارية على عينة من الأشخاص تمثل المجموعة البشرية المراد اختبارها لاحقاً.

تحضير فقرات الاختبار

يبحث الفصل الخامس من هذا الكتاب إجراءات وضع فقرات الاختبار الشفهية واختبارات الأداء بينما تتناول الفصول ٦ -٨ فقرات الاختبار المناسبة للاختبارات بشكل عام ولاختبارات الاستعداد بشكل خاص. وتنطبق العديد من التوصيات المتعلقة بوضع فقرات اختبارات التحصيل على أنواع أخرى من المقاييس الإدراكية.

ويمكن تقسيم فقرات الاختبار المستخدمة في تلك التي تستخدم فيها الورقة والقلم إلى الفئات التالية: الاستذكار بالمقابل مع الإدراك، أو الاستجابة المركبة بالمقابل مع التعرف على الأشياء، أو الفقرات المقالية مقابل الأسئلة الموضوعية. وتتطلب الأسئلة المقالية من الممتحن استذكار المعلومات المطلوبة من الذاكرة أو تزويد القائم على الاختبار بهذه المعلومات أو تركيب هذه المعلومات، في حين تتطلب الإجابة على الأسئلة الموضوعية عمليات إدراك الإجابات الصحيحة أو التعرف عليها أو اختيارها. وقد ازدادت شعبية فقرات الاختبارات الموضوعية بسبب سهولة تصحيحها وكونها أكثر موضوعية من الأسئلة المقالية. ومن الضروري لفت نظر القارئ هنا إلى أن فقرات الاختبار الموضوعية والمكتوبة بشكل محكم تعطي الممتحن فكرة أكثر وضوحاً من الأسئلة المقالية عن المهمة التي يجب إنجازها لاجتياز الاختبار بنجاح، إلا أن الموضوعية الفعلية لهذا النوع من الأسئلة لا تكمن في الأسئلة بحد ذاتها إنما في عملية وضع الدرجات. وهذا

يعني أن مصححَين مختلفين سوف يضعان الدرجات نفسها للاختبار نفسه، على حين يصعب حدوث هذا في حالة الأسئلة المقالية.

فقرات الاختبار المقالية

على الرغم من كون الاختبارات الموضوعية مقاييس فعالة للذاكرة الإدراكية إلا أن هذه الاختبارات تنقصها القدرة على قياس العمليات الإدراكية المعقدة كالقدرة على التنظيم والسرد والتواصل. تقيس الأسئلة المقالية هذه العمليات كما أنها تتطلب وقتاً أقل للإعداد ولا تترك للمجيبين عليها مجالاً للتخمين. إلا أن هذه الأسئلة، ومع الأسف، غالباً ما تكون أسئلة عامة جداً مما يجعل معناها يختلف من شخصٍ إلى آخر. وعلاوةً على ذلك لا يمكن للممتحن الإجابة إلا على عدد محدود جداً من الأسئلة المقالية ضمن فترة الاختبار، مما لا يسمح بتغطية المعرفة العلمية التي سيتم تقييم الممتحنين على أساسها. أما الممتحنون الذين لا يملكون فكرة واضحة عن المادة العلمية موضع الاختبار فغالباً ما يكون تقييمهم مخادعاً. وأخيراً تبقى المشكلة الكبرى في الاختبارات المقالية كون هذه الاختبارات تتطلب وقت تصحيح طويل جداً مقارنة بالأسئلة الموضوعية.

وتتوفر عدد من التوصيات التي يمكن أن تجعل الاختبارات المقالية أكثر موضوعية وأكثر شمولاً:

١- تحديد هدف السؤال والتأكد من أن معناه شديد الوضوح مما يسمح للخبراء بالتوصل إلى القرارات ذاتها فيما يتعلق بكون إجابةٍ ما إجابةً جيدة أو سيئة.

٢- استخدام عدد أكبر من الأسئلة المقالية التي تتطلب إجابات قصيرة بدلاً من عدد أقل من الأسئلة ذات الإجابات الطويلة.

٣- تحديد الطول المطلوب للإجابة (صفحة أو نصف صفحة... إلخ).

٤- الطلب من الممتحنين الإجابة على جميع فقرات الاختبار والطلب منهم وضع إجابة كل فقرة على ورقة منفصلة.

فقرات الإجابات القصيرة

تقع الفقرات الاختبارية التي تطلب من الممتحن إجابات قصيرة كملء الفراغات أو الإجابة على سؤال بعدد من الكلمات أو العبارات بين الأسئلة ذات الإجابات المقالية الطويلةوأسئلة التعرف (كأسئلة الخيارات المتعددة وأسئلة الصواب أو الخطأ وأسئلة المطابقة). وعلى الرغم من سهولة وضع أسئلة الإجابات القصيرة، تعد هذه الأسئلة محدودة من حيث تقييمها لمعرفة المصطلحات أو الحقائق كما يمكن أن تقود إلى أخذ الإجابات من النص مباشرةً ويمكن أن تحتمل إجابتين أو أكثر مما يجعل هذه النوع من الأسئلة أقل موضوعية مما هو متوقع. ويجب أن تصاغ أسئلة الإجابات القصيرة بشكل لا يحتمل الغموض أو الالتباس، وتعد الأسئلة المباشرة أفضل بكثير من العبارات الناقصة، ويفضل وضع الفراغ في نهاية العبارة وليس في منتصفها. ويجب تفادي وضع الفراغات كلها في قسم محدد من السؤال ومن الأفضل أن تكون جميع الفراغات من الطول ذاته. وفي حال كون الإجابات المطلوبة إجابات عددية يجب أن يتم تحديد الوحدة المطلوبة للإجابة.

يحتوي برنامج F-1 الموجود على القرص المرن المرافق لهذا الكتاب على نموذج اختبارملء فراغات للعمليات الحسابية البسيطة (الجمع والطرح والضرب والتقسيم ورفع الأعداد المكونة من آحاد أو آحاد وعشرات إلى أس معي). ومن ميزات هذا البرنامج كون العددين الذين يولدهما البرنامج للمسائل المختلفة عددين عشوائيين مما يعني حصول كل ممتحن على مجموعة مختلفة من المسائل.

ويعد نمط أسئلة ملء الفراغات ضمن المقاطع النصية (أسئلة كلوز cloze) نمطاً مناسباً جداً لتقييم مهارات القراءة. إذ تعتمد هذه الطريقة على حذف عدد معين من الكلمات من نصٍ ما ويطلب من الممتحن أن يجد كلمات مناسبة لوضعها مكان ما تم حذفه. وتعد درجة قدرة الممتحن على فهم النص وملء الفراغات بكلمات مناسبة مقياساً لقدرات الممتحن في مجال القراءة. ومن الأمثلة على أدوات القياس هذه كتاب اختبارات النصوص الناقصة (هودر وستوفتون) التي تحتوي عادةً على ٥٠ فراغاً. وتفيد هذه الاختبارات ليس في قياس قدرة الأفراد على القراءة فحسب بل تفيد المعلومات عن مدى تمكن مجموعة من الأفراد من ملء فراغات النص بنجاح في تحديد ما إذا كان النص بحد ذاته مقروءاً (يسهل استيعابه) أم لا.

أسئلة الخطأ والصواب

على الرغم من أن أسئلة الخطأ والصواب هي أسئلة سهلة الإعداد إلا أنها غالباً ما تُستخدم بشكلٍ خاطئ. وتُمكن سهولة وسرعة قراءة أسئلة الصواب والخطأ الممتحنين من الإجابة على ضعف عدد أسئلة الخيارات المتعددة في نفس المدة الزمنية. إلا أن هذا النوع من الأسئلة، شأنه كشأن الأسئلة ذات الإجابات القصيرة، يتيح للممتحن فرصة الحصول على الإجابات مباشرة من النص كما أنه غالباً ما تكون الأسئلة سطحية وتتعامل مع أمور بدهية. ويتوجب على مصممي الاختبارات من نوع اختبارات الخطأ والصواب، تفادي استعمال أدوات الدلالة المحدِّدة Specific determiner مثل: غالباً أو أحياناً أو عادةً التي ترتبط عادةً بالعبارات الصحيحة وكلمات مثل: كل ودائماً وأبداً وفقط التي ترتبط بالعبارات الخاطئة عند وضع هذا النوع من الأسئلة. يمكن أن يقود استخدام أدوات الدلالة المحددة التي يمكن أن تضفي هالة من الصدق على المقولات الكاذبة واستخدام المقولات التي تحتوي على منطقٍ سطحي أو اعتقادات

خاطئة شائعة أو فيها صدى الحقيقة بعض الأشخاص غير المتمكنين إلى اختيار بدائل الإجابة الخاطئة. ويجب تفادي كذلك الجمل التي تحتوي على تصريحات عامة جداً أو التي تستخدم النفي (والنفي المزدوج بشكلٍ خاص) أو الجمل الشديدة التعقيد والغموض أو الجمل الطويلة المعقدة أو التي تحتوي عدداً من الأفكار أو آراءاً شخصية (إلا في حال ذكر المصدر الذي تم أخذ الرأي منه). وأخيراً يجب أن تتساوى أسئلة الصواب والخطأ في طولها وعددها. وتحتوي برامج F-3 وF-4 على اختبارين مؤلفين من ٥٠ سؤال صواب و خطأ يتناول الأول منهما موضوع الشيخوخة والتقدم في العمر بينما يتناول الثاني موضوع الموت. ويمكن إجراء الاختبارين وتصحيحهما عن طريق البرنامجين F-3 وF-4.

أسئلة المطابقة

تعد فقرات الاختبار التي تعتمد على وضع عمودين يحتوي الأول منهما على عدد من الفرضيات Premises (أو المحفزّات) بينما يضم الآخر الإجابات المناسبة والتي يطلب من الممتحن أن يتعرف عليها نوعاً من فقرات الاختبار القليلة الاستخدام والتي يندر وجودها في الاختبارات القياسية. وتعد فقرات الاختبار التي تعتمد على إعادة الترتيب نوعاً من فقرات المطابقة التي تتطلب من الممتحن ترتيب احتمالات الإجابة ضمن مجموعات مرتبة. ويراعى عند تصميم فقرات المطابقة ترقيم الفرضيات تسلسلياً وبإستخدام الأعداد وترقيم اختيارات الإجابة بحروف متسلسلة. ويعد عدد من الفرضيات يتراوح بين ٦ إلى ١٥ عدداً معقولاً يرافق كل فرضية منها اختياران أو ثلاثة. ومن الضروري وضع الفقرة الاختبارية بكاملها على صفحة واحدة وتحديد الأساس الذي يتم الاستناد إليه عند المطابقة بين الفرضيات والإجابات. تعد فقرات المطابقة، بالرغم من كونها مقاييس لمعرفة الحقائق المحددة (راجع شاها، ١٩٨٤، لرأي مغاير)، فقرات اختبارية سهلة التصميم وتغطي المادة العلمية المراد اختبارها بشكل جيد بالمقارنة

مع غيرها من أنواع الفقرات الاختبارية. إن تحديد خيارات الإجابة بنفس النوع من المعلومات (أسماء، تواريخ، أماكن، إلخ) يعدّ محدداً رئيساً في اختبارات مطابقة الفقرات.

يمكن استخدام برنامج B-4 لإجراء وتصحيح اختبارات المطابقة وإعادة الترتيب، ويحتوي الملف "match.1" على العمود الأول أو الفرضيات بينما يحتوي الملف "match.2" على العمود الثاني أو عمود الإجابات، أما مفتاح الإجابات فيوجد في الملف "match.3" تحتوي الملفات أعلاه على الفرضيات والإجابات ومفتاح الإجابات لسؤال مطابقة يتناول تاريخ التقييم النفسي ويمكن استبدال محتويات هذا السؤال بأي معطيات أخرى للحصول على أسئلة مطابقة أو إعادة ترتيب جديدة.

أسئلة الخيارات المتعددة

تعد نماذج الاختيارات المتعددة أكثر صور الاختبارات الموضوعية تعدداً للاستعمال، وهي الشكل الأكثر استخداماً في الاختبارات الموحدة القياسية. تقيس هذه الفقرات المهارات المعقدة والبسيطة ويتم تصحيح هذه الفقرات بشكل أسهل وأسرع بكثير من الأسئلة المقالية. وبالرغم من إحتمالية إعطاء الممتحن للإجابة الصحيحة عن طريق التخمين، تبقى هذه الاحتمالية أقل منها في حالة أسئلة الصواب والخطأ. ويمكن الحصول على معلومات تشخيصية مفيدة عن طريق تحليل الإجابة للاختيارات الخاطئة (المشتّتات، الموهات). لا يخلو هذا النوع من فقرات الاختبار من العلل، ولعل من أهمها كون عملية وضع هذا النوع من الأسئلة تتطلب وقتاً وجهداً كبيرين. كما أن هذا النوع من الأسئلة يركز على مهارات التعرف بدلاً من مهارات التنظيم ويتطلب وقتاً أطول للإجابة من أسئلة الصواب والخطأ، كما يمكن أن لا يغطي هذا النوع من الأسئلة المادة العلمية موضع الاختبار وأن يعاقب الأفراد الذين يفكرون بعمق (راجع هوفمان، ١٩٦٢).

يحتوي الصندوق ٣-١ على قائمة توصيات مستقاة من أفضل الممارسات في مجال وضع أسئلة الخيارات المتعددة. إن برامج الحاسب التي تساعد على إعداد فقرات الاختبار (مثال: COURSEWRITER، PLANIT، TUTOR)، ومواصفات المواد والمقررات الدراسية الخوارزمية التي تساعد على توليد أسئلة اختبارات، وطرق كتابة فقرات الاختبار التي تستند إلى المهارات اللغوية ومعرفة المفاهيم، تَعِدُ بإجراءات أكثر فاعلية ودقة لإعداد أسئلة الخيارات المتعددة (هيرمان، ١٩٩٤).

وبرغم وجود قواعد وخوارزميات Algorithms لكتابة أسئلة الخيارات المتعددة، تبقى عملية الكتابة هذه نوعاً من الفن والعلم في آنٍ واحد. ويمكن تعقيد مهمة الممتحن باستخدام العبارات التالية: "كل ما ذكر أعلاه" و"لا شيء مما ذكر أعلاه" و"شيئان مما ذكر أعلاه" و "كل ما ذكر أعلاه إلا شيء واحد" و"واحد مما يلي." كما يمكن أن تكون جميع الإجابات صحيحة إلا أن على الممتحن اختيار الإجابة الأفضل أو الأكثر صحة. ومن المقترحات الأخرى لجعل الأسئلة متعددة الخيارات أكثر صعوبة: الطلب من الممتحن أن يتعرف على الإجابات الصحيحة من ضمن عدد متغير من الإجابات أو إجراء تحسينات (تعديلات) على الإجابات المعطاة أو التعرف على النص الصحيح لمسألة رياضية ما.

يحتوي الصندوق ٣-٢ على مقترحات أخرى لجعل اختيار الإجابات الصحيحة في أسئلة الخيارات المتعددة عملية تتطلب تحليل وتفكير ولا تقتصر فقط على الحفظ عن طريق التكرار.

الصندوق ٣-١

توصيات لإعداد الأسئلة المتعددة الاختيارات

١- اجعل جذر السؤال إما جملة غير مكتملة أو سؤال (وهذا الأخير هو الخيار الأفضل). وفي حال استخدام جملة غير مكتملة يجب وضع الفراغ في نهاية الجملة.

٢- تأكد من صياغة جذر السؤال صياغة واضحة ومتناسبة مع إمكانيات المتحنين اللغوية.

٣- لا تكثر من استخدام أسئلة الآراء وفي حال استخدام هذا النوع من الأسئلة أشر بوضوح إلى مصدر الرأي.

٤- حاول أن يشتمل جذر السؤال على أكبر قدر ممكن من الفقرة الاختبارية بما فيه العبارات المتكررة التي تنطبق على جميع البدائل (الاختيارات).

٥- اجعل الاختيارات كلها متساوية الطول وصحيحة من حيث ارتباطها النحوي مع جذر السؤال ومعقولة لمن لا يعرف الإجابة على السؤال.

٦- حدد، عند صياغة المموهات، الأسباب التي يمكن على أساسها أن يختار المتحنون الذين يجهلون الإجابة الصحيحة هذه المموهات.

٧- الأخطاء الشائعة التي تعد صحيحةً جزئياً هي نوع جيد من المموهات.

٨- تأكد من عدم وجود أية إشارات لغوية أو أي أدلة ضمن جذر السؤال تشير إلى الإجابة الصحيحة.

٩- على الرغم من أن معظم أسئلة الخيارات المتعددة تضم أربعة أو خمسة خيارات، يمكن الإكتفاء بخيارين أو ثلاثة وخصوصاً في المراحل الدراسية الدنيا.

obeikandl.com

١٠- رتب الخيارات التي تتطلب ترتيباً بطبيعتها (الأعمار أو التواريخ) تبعاً لترتيبها الطبيعي أو بشكل عشوائي أو وفق الترتيب الهجائي.

١١- حاول تفادي (أو الإقلال من) استخدام أدوات النفي (ليس) أو أدوات الدلالة المحددة (دائماً وأبداً) و جمل مثل: "كل ما ورد أعلاه" أو "لا شيء مما ورد أعلاه" أو "أكثر من شيء مما ورد أعلاه".

١٢- لا تستخدم الأسئلة المتداخلة (التي تقود الإجابة على واحد منها إلى الإجابة على الثاني).

١٣- رتب الخيارات على صورة فقرات (تراكمات) وليس على هيئة ازدواجات الواحد خلف الآخر).

١٤- استخدم الأعداد لترقيم الأسئلة والأحرف لترقيم الخيارات.

١٥- تأكد من مراجعة وتدقيق الفقرات من قبل مصمم اختبارات ذو خبرة واسعة قبل استخدامها.

الصندوق ٣-٢

أمثلة عن أسئلة الاختيارات المتعددة المعقدة

١- التصنيف. يطلب من الممتحن أن يصنف الأشياء أو الأشخاص أو الحالات إلى عدد من الفئات التي تم ذكرها في جذر السؤال. مثال:

يمكن وصف جان بياجيه على أنه عالم نفس ----------.

أ. سريري ب. سايكومتري ج.تطوري د. اجتماعي

٢- حالات إذا - فإن. يحدد الممتحن النتائج الصحيحة لحالة أو عدد من الحالات الموجودة. مثال:

إذا ازداد التباين الحقيقي للاختبار وبقي تباين الخطأ على ما هو عليه، فأي من الأمور التالية سوف

يحدث:

أ- سيزداد ثبات الاختبار.

ب- سيقل ثبات الاختبار.

ج- سينخفض التباين الملاحظ.

د- لن يطرأ أي تغير على الثبات أو التباين الملاحظ.

٣- الشروط المتعددة. يستخدم الممتحن الشروط الموجودة في جذر السؤال (اثنان أو أكثر) ليتوصل إلى استنتاج معين. مثال:

إذا كانت درجة ماري الأولية (Row score) في اختبار ما ٦٠ وكان الوسط الحسابي للاختبار ٥٩ والإنحراف المعياري ٢، فما هو معدل درجات ماري z:

أ. ٢,٠٠- ب. ٥٠. - ج. ٥٠. - د. ٢,٠٠

٤- أسئلة الصواب والخطأ المتعددة الاختيارات. يقرر الممتحن صحة شرط واحد أو شرطين أو أكثر أو عدم صحة أي من الشروط المذكورة في جذر السؤال. مثال:

هل من الصحيح أن ١) ألفرد بينيه كان أب اختبارات الذكاء وأنه ٢) نشر أول اختبارات الذكاء التي وضعها في عام ١٩١٦؟

أ. ١ و ٢ صحيحان ب. ١ صحيح لكن ٢ غير صحيح

ج. ٢ صحيح ولكن ١ غير صحيح د. ١ و ٢ خطأ

obeikandl.com

٥- الإختلاف. يشير الممتحن إلى الخيار الذي لا ينتمي لباقي المجموعة. مثال:

أي من الأسماء التالية هو المختلف؟

أ. ألفرد أدلر. ب. سيغموند فرويد. ج. كارل يونغ. د. كارل روجرز

٦- العلاقات والروابط. يحدد الممتحن العلاقة بين المفاهيم ١ و٢ ثم يختار من بين الخيارات (أ، ب، ج، د... إلخ) المفهوم الذي يرتبط مع ٢ بعلاقة مشابهة للعلاقة بين ١ و٢ مثال:

الوسط الحسابي بالنسبة للانحراف المعياري هو كالوسيط بالنسبة إلى:

أ. الانحراف الوسطي. ب. المدى الشامل.

ج. شبه المدى الربيعي Semi - inter quartile range. د. التباين

وقد أدت صعوبة وضع أربعة أو خمسة اختيارات متساوية الطول وقادرة على جذب الممتحن الذي يجهل جواب السؤال إلى اتجاه واضعي الاختبارات إلى النموذج ذو الخيارين. ويمثل برنامجا F-2 وF-5 الموجودان على القرص المرن المرفق بهذا الكتاب اختبارات تحتمل فقراتها خيارين فقط. يضم القسم الأول اختباراً في تهجئة ٢٨ كلمة غالباً ما يقع الناس في الخطأ عند تهجئتها ويضم القسم الثاني اختباراً في تهجئة ٨ أنواع من الجموع (المفردات في حالة الجمع) بينما يحتوي القسم الثالث على اختبار ٢٠ كلمة لها معانٍ مختلفة. وتوفر هذه البرامج الدرجات الخام والنسبة المئوية للإجابات الصحيحة للاختبار بكامله ولكل قسم على حدة. كما يُستخدم برنامج F-5 لإجراء وتصحيح اختبار مؤلف من ٥٥ فقرة عن المصطلحات المستخدمة في المقررات التمهيدية لمادة علم النفس. يتيح هذا البرنامج للممتحنين فرصة معرفة درجاتهم أثناء الاختبار ويسمح لهم بتغيير إجاباتهم ومن ثم مقارنة إجاباتهم مع الإجابات الصحيحة بعد انتهائهم من الاختبار. ويوفر البرنامج معلومات عن عدد الإجابات الصحيحة

ومعدل الدرجات المصحح من أثر التخمين والنسب المئوية للدرجات النهائية المحتملة.

ترتيب فقرات الاختبار

يعد ترتيب فقرات الاختبار التي تنتمي إلى نوع واحد والتي تتناول موضوعاً واحداً إلى جانب بعضها أمراً من مقتضيات السهولة وحسن التقديم، وعليه ترتب الأسئلة المتعددة الخيارات في قسم خاص وكذلك أسئلة الصح أو الخطأ وهكذا دواليك. وغالباً ما توضع الأسئلة المقالية التي تستغرق الإجابة عليها وقتاً طويلاً والتي تحتاج إلى طريقة تفكير مختلفة من قبل الممتحن في نهاية الاختبار، إلا أنه لا توجد قاعدة ثابتة لهذا الترتيب. وإن عدم تشتت الممتحن وتقليص عدد التحولات الإدراكية التي تطرأ عليه خلال فترة الاختبار هي الأهداف الرئيسة لتقسيم الاختبار إلى أقسام يضم كل منها أسئلة متشابهة. كما ترتب الفقرات الاختبارية عادةً من الأسهل إلى الأصعب لسبب بسيط وهو أن الأسئلة السهلة ترفع من مستوى ثقة الممتحن بنفسه وتشجعه على محاولة الإجابة على الأسئلة الأكثر صعوبة. إلا أن الأبحاث في هذا المجال أثبتت أن ترتيب فقرات الاختبار من الأسهل إلى الأصعب ليس له أي تأثير على الدرجات النهائية (أليسون، ١٩٨٤؛ كليمكو، ١٩٨٤) باستثناء نوعين من الاختبارات وهما الاختبارات الصعبة جداً واختبارات السرعة التي أثبت فيها أن ترتيب الأسئلة من الأسهل إلى الأصعب ساهم في رفع الدرجات بعض الشيء (غرين، ١٩٨٤؛ بلايك وأنسورج وباركر ولاوري، ١٩٨٢؛ سافيتز، ١٩٨٥).

وترتب فقرات بعض اختبارات الذكاء الجماعية على شكل مجموعة من الاختبارات الفرعية المنفصلة أو في اختبارات أخرى فإنها ترتب على شكل كتيبات عادية أو لولبية ترتب فيها الفقرات الاختبارية التي تقيس العمليات المختلفة التي تساهم في تكوين القدرة العقلية للأفراد من الأسهل إلى الأصعب.

وفي هذا النوع من الكتيبات تظهر الأسئلة التي تتناول موضوعاً معيناً بشكلٍ متناوب مع أسئلة من أنواع أخرى ولكن على نفس القدر من الصعوبة وترتب لولبياً من الأسهل إلى الأصعب.

تعليمات الاختبار

تكتب تعليمات الاختبارات الموضوعية أو المقالية التي تجرى في الوقت ذاته لمجموعة من الأشخاص على ورقة توضع في بداية الاختبار بينما توضع التعليمات الخاصة بأقسام الاختبارات المتعددة الفروع في بداية كل قسم. ويفضل طباعة هذه التعليمات بالحروف السوداء العريضة لتقليل احتمالية عدم انتباه الممتحنين إليها. وفي حال كون تعليمات الاختبار غير مألوفة أو معقدة، يفضل أن يقرأها المراقب أو القائم على الاختبار بصوتٍ عالٍ. أما في الاختبارات الفردية التي تسمح بالتواصل بين الممتحن والقائم على الاختبار، يعطي هذا الأخير التعليمات للممتحن شفهياً. ويشترط في تعليمات الاختبار(مقروءة كانت أم مطبوعة أم الاثنين معاً) أن تُطلع الممتحن على الأمور التالية:

- الهدف العام من الاختبار أو الفقرة الاختبارية.
- الوقت الزمني المتاح للممتحن.
- كيفية وضع الدرجات.
- ما إذا كان التخمين في الإجابة أمراً ينصح به (في حالة الاختبارات الموضوعية فقط).

تمثل التعليمات أدناه نموذجاً عن تعليمات الاختبارات:

اكتب اسمك في الزاوية اليمنى أعلى الصفحة ولا تكتب على ورقة الأسئلة. يهدف هذا الاختبار إلى قياس درجة معرفتك في مادة الإحصاء للمبتدئين. تتوفر أوراق مسودة يمكن استخدامها لإجراء العمليات الحسابية. يتألف الاختبار من ٦٠ سؤالاً ولديك ٩٠ دقيقة لإتمام الاختبار. حدد إجابتك على كل سؤال بملء

الفراغ المناسب على ورقة الإجابة تحت الحرف الذي يمثل الإجابة الصحيحة. سيكون معدل درجاتك الأولي هو مجموع إجاباتك الصحيحة. وعلى الرغم من أن التخمين الأهوج لا يفيد كثيراً إلا أنه من الحكمة أن تحاول استبعاد أي من الخيارات التي تعلم أنها خاطئة قبل اختيار الإجابة. يسمح لك وقت الاختبار بالإجابة على جميع الأسئلة ومراجعة إجاباتك. في حال انهاء الاختبار قبل انتهاء الوقت، الرجاء الجلوس بهدوء لحين انتهاء الآخرين من الإجابة.

يتوجب على القائم على الاختبار قراءة التعليمات بوضوح وبطء وبالشكل الذي تظهر عليه وبعد الانتهاء من القراءة يعطى المتحنون الفرصة لطرح الأسئلة التي يتعين على القائم على الاختبار الإجابة عليها (بشكل هادئ ومفيد) وإن كانت سخيفة أو مكررة.

أوراق الإجابة

لا ينصح باستخدام الاختبارات التي تتطلب أوراق إجابة منفصلة مع مجموعات الأطفال دون سن التاسعة أو العاشرة (أيراسيان وتراسي، ١٩٩٤) إذ من الأفضل الطلب من الأطفال دون هذا العمر كتابة الإجابات في هامش ورقة الأسئلة. أما في سنوات الدراسية الابتدائية العليا يسهّل استخدام أوراق إجابة منفصلة عمليتي وضع الدرجات وتحليل الإجابات. وفي هذا النوع من الاختبارات التي تصحح عن طريق الآلات، يطلب من المتحنين ملء الدائرة المرقمة (بالأحرف أو الأرقام) والموجودة إلى جانب السؤال.

إجراء الاختبار

تتأثر درجات الاختبارات النفسية ليس بمحتوى الاختبارات فحسب بل بالطريقة التي يتم فيها إجراء الاختبار. وبالرغم من الجهود المضنية التي تبذل لوضع اختبارات مناسبة من حيث الشكل والمضمون إلا أن هذا الأمر لا يضمن

وجود الدافع لدى جميع المتحنين للإجابة على أسئلة الاختبار بشكل صادق وملتزم كما لا يضمن فهم هؤلاء المتحنين لما هو مطلوب منهم. علاوة على ذلك، من السهل أن يصاب المتحنون بالإرهاق أو أن يتشتت انتباههم أو، لأسباب أخرى، يمكن أن يعاني المتحنون من فقدان التركيز أو فقدان الرغبة في إتمام الاختبار. وتساهم هذه العوامل مجتمعة في التأثير على معدلات درجات هؤلاء الأفراد مما يعني أن الاختبار فشل في رسم صورة دقيقة عن إمكانياتهم.

وتتنوع الإجراءات التي تُتخذ لضمان أن نتائج اختبارات الاستعداد تقيس بشكل فعال ودقيق المتغيرات التي أعدت لقياسها بتنوع الاختبارات بحد ذاتها. إذ تتطلب الاختبارات الفردية درجة أعلى من الاهتمام من قبل القائم على الاختبار مما تتطلبه الاختبارات الجماعية. وتختلف الإجراءات المرتبطة باختبارات الاستعداد عن تلك المرتبطة باختبارات الشخصية. وعلى إختلاف الاختبارات (فردية أم جماعية محددة بوقت معين أم وقتها مفتوح)، وبغض النظر عن الوقت الذي يجرى فيه الاختبار وعن الخلفية الحضارية أو العرقية للممتحن، يتوجب على القائم على الاختبار أن يكون على درجة عالية من الاستعداد. ويجب، لضمان أن تكون نتائج الاختبارات مقاييس صادقة لما يهدف الاختبار إلى قياسه، إعداد وإجراء الاختبار بشكل حذر والتأكد من أن القائم على الاختبار لديه الخبرة والتدريب اللازمين لإجراء الاختبار وأن الجو في قاعة الاختبار مناسب ومريح.

يتطلب إجراء اختبارات الذكاء واختبارات الشخصية الفردية أن يكون القائم على الاختبار لديه المؤهلات المطلوبة والتي تكون عادة أكثر صرامة من المؤهلات المطلوبة لإجراء الاختبارات الجماعية. ومن الضروري أن يكون القائمون على الاختبار على دراية تامة بحدود إمكاناتهم وأن يطلبوا المساعدة من الأفراد المؤهلين إذا نقصتهم الخبرة أو المعرفة. وفي حال اختبار القاصرين يجب

الحصول على موافقة ولي الأمر أو الوصي القانوني قبل إجراء الاختبار. ويفيد توقيع إستمارة مثل تلك الموجودة في الصندوق ١-١ في حل هذه المشكلة. وهنالك إستثناءات لهذه القاعدة إذ لا يطلب توقيع إستمارة الموافقة على الخضوع للاختبار في حال كون الاختبار ضرورة قانونية منصوص عليها أو في حال كون الاختبار جزءاً من النشاطات المدرسية الاعتيادية أو أن تكون الموافقة على الخضوع للاختبار متضمنة في عقد ما كطلب التوظيف أو طلب الإنتساب إلى كلية أو مؤسسة تعليمية ما.

وعلى الرغم من أن الأشخاص غالباً ما يرقون إلى مستوى المناسبة ويكون أداؤهم حسناً حتى تحت أحلك الظروف فإن الإضاءة وتوزيع المقاعد والتهوية والحرارة ومستوى الضجيج كلها تساعد على ارتفاع مستويات الأداء في الاختبارات. ويقوم القائم على الاختبار بالتعرف على قاعة الاختبار (المقاعد، الإضاءة، موقع النوافذ ومظّلمات الحرارة... إلخ) والتعرف على الحمامات والمخارج القريبة ثم يجري تعديلات إذا لزم الأمر. ويفيد وضع لوحة مكتوب عليها "امتحانات! يرجى الهدوء" في تخفيف المقاطعات والضجيج. وفي الاختبارات المدرسية ينصح بعدم إجراء الاختبارات بعد الغداء أو قبل العطلة أو غيرها من الأوقات التي يكون الطلاب فيها متعبين أو منفعلين.

ضمان سلامة الاختبارات والغش

بالإضافة إلى حاجتها لموظفين خبراء، يتطلب إجراء الاختبارات التي يسهل الغش فيها مثل اختبار التقييم المدرسي (SAT) واختبارسجل الخريجين (GRE) إلى رقابة مشددة أثناء فترة الاختبار. ترقم أوراق الاختبار بالتسلسل مما يسهل عملية إحصاء هذه الأوراق قبل وبعد الاختبار كما يطلب من الممتحنين إبراز هوياتهم الشخصية أو أي نوع آخر من بطاقات الهوية ولا يسمح لهم بمغادرة قاعة الاختبار إلا بعد انتهاء الاختبار.

وبغض النظر عن نوعية الاختبار، يبقى خطر الغش واقعاً تعاني منه عملية الاختبار، وغالباً ما يحصل الغش عندما يتمكن أفراد غير مخولين من الحصول على نسخ من أوراق الأسئلة أو غيرها من المعلومات عن فقرات الاختبار. ومن أكثر أشكال الغش شيوعاً إحضار أوراق صغيرة تحتوي على إجابات إلى قاعة الامتحانات أو نسخ الإجابات من ممتحنين آخرين. ومن الجدير بالذكر أنه لا توجد طريقة أكيدة لمعالجة هذه المشكلة ولكن ينصح بوضع مقاعد المتحنين على مسافات بعيدة (وترك مقعد واحد فارغ على الأقل) والطلب من المتحنين إفراغ محتويات جيوبهم ومقاعدهم وتكثيف عدد المراقبين في قاعة الامتحان والتأكد من أوراق الأسئلة قبل وبعد الاختبار. كما يساعد برنامج B-7 الموجود على القرص المرافق لهذا الكتاب في التعرف على عدد من حالات الغش في الاختبارات.

يمكن أن تحدث حالات الغش قبل أو بعد أو أثناء الاختبار فعلى سبيل المثال، إكتشفت منذ عدة سنوات حالة غش واسعة النطاق في كاليفورنيا إذ كان المدرسون في بعض مدارس الولاية يمسحون الإجابات الخاطئة التي اختاراها الطلاب في اختبار برنامج كاليفورنيا للتحصيل (CAP) ويضعون الإجابات الصحيحة مكانها وذلك بهدف تحسين مظهر مدارسهم وطلابهم. كما يمكن أن تحدث حالات غش بعد انتهاء الاختبار إذ يمكن إضافة نقاط إلى معدلات درجات طالب أو موظف معين بهدف جعل هذا الفرد يبدو أكثر كفاءة لأي سبب من الأسباب.

تعليمات الاختبار وإجراءاته

من المهم التقيد بتعليمات الاختبارات القياسية وخاصةً في الحالات التي سيتم تحويل درجات الاختبار فيها إلى معايير عامة مستمدة من مجموعة الأفراد الذين تم اختبارهم وفق هذه التعليمات. إلا أنه، وفي حالات الاختبارات

السريرية، لا يمكن دائماً التقيد الدقيق بتعليمات الاختبار ففي حالة اختبار شخص معاق جسدياً أو اختبار حدود الإمكانيات الجسدية أو العقلية لفردٍ ما يفضل عدم التقيد بالتعليمات النموذجية.

يتعين على القائم على الاختبار في حالة الاختبارات الجماعية والفردية التعريف بنفسه ثم التحدث مع المتحنين للتعرف عليهم بشكل سريع. وفي حالة الاختبارات الفردية تتطلب عملية التعرف والتواصل هذه جهداً أكبر ووقتاً أطول إذ يتعين على القائم على الاختبار الوصول إلى درجة جيدة من التواصل مع المتحن بهدف إعطاء المتحن شعوراً بالطمانينة مما يمكن أن يزيل مخاوف المتحن ويساعده على تحقيق مستوى أداء أعلى في الاختبار. وغالباً ما تساعد خبرة القائم على الاختبار في التعامل مع أفراد من نفس العمر والجنس والعرق والخلفية الحضارية واللغوية الذي ينتمي إليه المتحنون على مساعدة المتحنين على الوصول إلى مستوى أدائهم الأمثل. ويتطلب اختبار أفراد متقدمين في العمر أو أطفال صغار أو أفراد لديهم إعاقات معينة أو أفراد ينتمون إلى حضارة مختلفة الخروج عن إجراءات الاختبار وتعليماته النموذجية، فيمكن مثلاً تغيير صياغة الأسئلة لجعلها أكثر بساطة أو زيادةوقت الاختبار أو تقصيره أو تشجيع المتحنين بشكل دوري ومتواتر أو توفير أدوات مساعدة على التحرك أو الإحساس. ويجب مراعاة في حال تعديل إجراءات الاختبار لمجموعة معينة من الأفراد أنه لا يمكن مقارنة أداء هذه المجموعةمع أداء عينة القياس. ومن الحكمة أحياناً عند اختبار الأطفال المعاقين أن يعطى هؤلاء الأطفال اختبارات أدنى من مستوى عمرهم أو سنتهم الدراسية (وهذا ما يسمى الاختبار الخارج عن المستوى). وتتوفر في هذه الحالات معايير خاصة للاختبار الخارج عن المستوى تقيم أداء المتحنين في الاختبارات القياسية.

البراعة في الاختبارات

يُعتبر بعض الأفراد أقدر من غيرهم على تحقيق أداء جيد في الاختبارات ويعزى هذا إلى عدد من العوامل كمستوى ذكاء أعلى أو مستوى قلق أدنى أو بسبب كون الطالب أكثر جاهزية أو أبرع في الاختبارات لسبب مجهول. وغالباً ما يرتفع مستوى القلق لدى الأفراد في الاختبارات المحدودة الوقت، ومما لا شك فيه أن هذا القلق يزيد عندما يكون الممتحن على درجة غير مقبولة من الجاهزية للاختبار. وعلى الرغم من أن القلق وعدم الجاهزية يؤثران على معدلات درجات الاختبارات، يلعب عامل يصعب تعريفه هو عامل "البراعة في الاختبارات" دوراً في التأثير على الدرجات. ويمكن لشخص بارع في الاختبارات وجاهل بمحتوى الاختبار أن يرفع مستوى أدائه عن طريق الاهتمام بإجاباته (إيفانز، ١٩٨٤). ويعد بعض الممتحَنون أكثر قدرة على التعرف على الإجابات الخاطئة في الأسئلة متعددة الخيارات حيث يمكنهم ملاحظة أن الإجابات الخاطئة غالباً ما تصاغ بطريقة مختلفة عن الإجابات الصحيحة (تكون الإجابات الخاطئة عامة جداً أو محددة جداً أو تحتوي على أخطاء نحوية). ويلاحظ هؤلاء الأفراد كذلك كون الإجابات الصحيحة دقيقة التعبير وتحتوي على دلائل نحوية أو أنها مرتبطة بإجابات أو فقرات اختبارية أخرى. وتتوفر أدلة على كون فقرات الاختبار اللغوية أكثر تأثراً بعامل "البراعة في الاختبارات" وعلى كون الذكور أكثر "براعة" من الإناث (برستون، ١٩٦٤؛ راولي، ١٩٧٤). هذا ويمكن تدريب الأفراد لزيادة براعتهم في الاختبارات (وهذا ما يوضحه الصندوق ٣-٣).

الاختبارات المحوسبة

توضح برامج الكمبيوتر B-2 وبرامج الفئات F وG وH الموجودة على القرص المرافق لهذا الكتاب سهولة إجراء ووضع درجات الاختبارات الموضوعية بوساطة الحاسب الآلي. وتعد برامج الاختبارات التكيفية المحوسبة

(CAT) Computer - based adaptive testing والتي تحدد الأسئلة اللاحقة فيها بأداء الممتحَن في الأسئلة السابقة نوعاً مطوراً من برامج الاختبارات المحوسبة السهلة. ويقوم برنامج الكمبيوتر باختيار السؤال اللاحق لممتحن ما بناءاً على إجاباته على الأسئلة السابقة، ويتم الاختيار من مجموعات من الأسئلة تمت معايرتها تبعاً لمستويات القدرة المختلفة وبناءً على نموذج الاستجابة للأسئلة IRT. يبدأ الممتحن بسؤال يعد مستوى صعوبته متوسطاً ومن الممكن أن يكون هذا السؤال مبنياً على عمر الممتحن أو درجاته في اختبارات سابقة أو غيرها من المعلومات الإضافية، وبعد فراغ الممتحن من الإجابة على السؤال الأول يحدد الكمبيوتر مستوى السؤال التالي الذي يكون من مستوى قدرة أعلى في حال إعطاء الممتحن إجابة صحيحة على السؤال أو من مستوى قدرة أدنى إذا أعطى الممتحن إجابة خاطئة، وتكرر هذه العملية طوال الاختبار.

يسمح هذا النمط من الاختبارات بالحصول على تقدير دقيق لقدرة الممتحن من خلال عدد من الأسئلة يشكل ٥٠٪ من عدد الأسئلة الموجودة في كتيبات الاختبار التقليدية.

الصندوق ٣-٣

اقتراحات لزيادة "البراعة" في الاختبارات

قبل الاختبار

١- اطلب من المدرس إطلاعك على أسئلة السنوات السابقة من نفس الاختبار والتي يُسمح لك الإطلاع عليها.

٢- إسأل الطلاب عن أنواع الاختبارات التي يأتي بها مدرس المقرر.

٣- لا تؤخر التحضير للاختبار إلى اليوم الذي يسبق الاختبار في حال كون يوم الاختبار قد أعلن عنه مسبقاً.

obeikandl.com

٤- ادرس بشكل يتناسب مع نوع الاختبار المعلن عنه (أسئلة متعددة الاختيارات، صواب أو خطأ، مقالات).

٥- في حال عدم الإعلان عن نوع محدد من الاختبارات ينصح بالتحضير للاختبارات المقالية التي تعتمد على الذاكرة.

٦- لا تحول عملية التحضير للاختبار إلى مناسبة اجتماعية إذ من الأفضل لك أن تنعزل وتدرس بمفردك.

٧- لا تجلس بشكل مريح جداً (كالاستلقاء على السرير) إذ قد يسيء جسمك فهم هذه الراحة ومن الجائز أن تشعر بالنعاس.

٨- حاول أن تُعد أسئلة اختبار من المعلومات التي تدرسها وحاول أن تكون هذه الأسئلة من نوع الخيارات المتعددة إذا كان الاختبار يعتمد على هذا النوع من الأسئلة، أو أسئلة مقالية إذا كان الاختبار مقالياً.

٩- حاول أن تطبق طريقة الدراسة المعروفة بـ Q3R (خذ فكرة عامة عن المادة أو المقرر، ضع أسئلة، اقرأ، استذكر، راجع). اقرأ أولاً لتكوّن فكرة عامة عن المادة العلمية ثم إسأل نفسك بعض الأسئلة حول هذه المادة ثم إقرأ المادة بتمعن يمكّنك من تذكر المعلومات ثم كرر المعلومات لتثبت في ذاكرتك وأخيراً راجع المعلومات قبل الاختبار.

أثناء الاختبار

١- اقرأ تعليمات الاختبار بتمعن قبل البدأ بالاختبار وفي حال وجود أي استفسارات حول أشياء لم تذكرها التعليمات كالزمن المتاح لإنهاء الاختبار أو أوزان فقرات الاختبار المختلفة أو ما إذا كان الممتحن يفقد درجات في حال التخمين الخطأ للإجابة أو غيرها، لا تتردد في سؤال المراقب أو القائم على الاختبارعنها.

obeikandl.com

٢- فكر ملياً بالإجابة أو استعمل المسودة قبل الإجابة على الأسئلة المقالية.

٣- وزع وقتك بالتساوي على الأسئلة ففي حالة اختبارات الإختيار من متعدد يجب أن تجيب على عدد مناسب من الفقرات يعادل الزمن الكلي للاختبار مقسماً على عدد الأسئلة.

٤- لا تترك أي سؤال دون إجابة إذا كان بإمكانك إستبعاد أي من الخيارات حتى في حال فقدان الدرجات في حالة التخمين الخطأ للإجابة.

٥- إترك الأسئلة الصعب التي لم تستطع الإجابة عليها وعد إليها بعد فراغك من الإجابة على الأسئلة الباقية. حاول أن لا تصاب بالذعر بسبب الأسئلة التي لا تعرف إجاباتها. ضع إشارة إلى جانب هذه الأسئلة وحاول أن تجيب عليها وإن لم تستطع فخمن عن علم وتفكير.

٦- في حال توفر الوقت لديك لمراجعة إجاباتك إفعل هذا ولا تكن في عجلة لتسليم ورقة الاختبار ومغادرة القاعة.

تتوفر عدد من الاختبارات تعتمد على مبدأ مشابه للاختبارات التكيفية كالاختبارات المتعددة المستويات والاختبارات ذات المستويات المرنة. وتعد الطبعة الرابعة من اختبار ستانفورد - بينيه واحدة من هذه الاختبارات المتعددة المستويات التي تعتمد على إجراء اختبار قصير لجميع المتحنين يطلق عليه اسم الاختبار "محدد المسار" وهو اختبار تحدد معدلات درجات المتحن فيه مستوى الاختبار الذي سيعطى للممتحن لاحقاً. وترتب الأسئلة في الاختبارات ذات المستويات المرنة تبعاً لمستوى الصعوبة من الأسهل إلى الأصعب ويطلب من جميع المتحنين البدء بالأسئلة التي تقع في المنتصف ثم تؤهل الإجابات الصحيحة المتحن للإجابة على الأسئلة الأكثر صعوبة بينما تقود الإجابات الخاطئة المتحن إلى الإجابة على الأسئلة الأكثر سهولة وتكرر هذه العملية طوال الاختبار (غرين، ١٩٩٢).

تتمتع الاختبارات التكيفية بمزايا عديدة ككونها أكثر دقة من أشكال الاختبار التقليدية وكونها تسمح بخلط احتمالات الإجابة لمنع تكرار الخيارات ذاتها كما تسمح باختبار كل ممتحن بناءاً على مستوى قدرته وبأن يحس الممتحن بالتحدي طوال الوقت. إلا أن أحد المشاكل المرتبطة بهذا النوع من الاختبارات هي عدم قدرة الممتحن على ترك سؤال والعودة إليه لاحقاً أو تغيير إجابته على السؤال، كما يتوجب على واضع الاختبار التأكد من وقوع السؤال برمته على صفحة واحدة تظهر على شاشة الكمبيوتر بشكل كامل. كما تعاني الاختبارات التكيفية من المشاكل التي تعاني منها نظرية تحليل الفقرات الاختبارية التي تعد أساس الاختبارات التكيفية. إلا أنه، وعلى كل حال، تعتبر الاختبارات التكيفية (CAT) اختبارات يستجيب إليها الممتحنون بشكل جيد (غرين، ١٩٩٢).

تصحيح الاختبارات وتفسير النتائج

يجب إتخاذ القرارات المتعلقة بتصحيح الاختبارات (إجراءات التصحيح وقواعده ومفتاح الإجابات) قبل إجراء الاختبار وخلال مرحلة التصميم. كما يجب تحديد الطريقة التي سيتم من خلالها التعبير عن نتائج الاختبارات عددياً، ويجب اعتماد طريقة لوضع الدرجات إما باليد أو عن طريق الحاسب. ومن الجدير بالذكر أن استخدام ورقة منفصلة للإجابة لا تعني بالضرورة استخدام الحاسب لوضع الدرجات وإن كان هذا الاستخدام يسهل من عملية التصحيح المحوسبة. وسواء كان التصحيح باليد أو بالحاسب الآلي، هل سيتم التصحيح على هيئة سلسلة من الأجزاء أو على الاختبار ككل أو كليهما؟ ومن الضروري إتخاذ قرار بشأن عدد الدرجات التي تعطى للإجابات الصحيحة ويتم الاختيار عادةً بين طريقة التصحيح البسيطة التي تمنح درجة واحدة للسؤال الصحيح أو طريقة التصحيح الأكثر تعقيداً التي تعطي أوزاناً مختلفة للإجابات المختلفة. كما يجب، في الاختبارات الموضوعية (أسئلة الاختيار المتعدد وأسئلة الصواب والخطأ)، أخذ مبدأ "التصحيح من أثر التخمين" بعين الاعتبار وهو مبدأ يتم على أساسه حذف

عدد من الدرجات الأولية للممتحنين الذين خمنوا خطأً عدداً من إجاباتهم. وفي الأسئلة المقالية يجب تحديد ما إذا كان يجب تقييم الإجابات بناءاً على عاملين منفصلين كالمحتوى والشكل (القواعد والترتيب... إلخ)، وإن كان الأمر كذلك هل يجب الجمع بين هذين العاملين بشكلٍ أو بآخر. والسؤال الأخير الذي يجدر طرحه هنا هو: هل سيتم تحويل النتائج الأولية للاختبارات إلى علامات نهائية (A، *A... إلخ) وهل يجب إرفاق ملاحظات مكتوبة مع ورقة النتائج؟

اختبارات الخيارات المتعددة والخطأ والصواب

تعتمد أبسط طرق وضع درجات الاختبارات المؤلفة من أسئلة الصواب والخطأ أو أسئلة الخيارات المتعددة على إعطاء درجة واحدة للإجابة الصحيحة ودرجة الصفر للإجابة الخاطئة مما يعني أن الحد الأدنى من الدرجات الممكن الحصول عليها هو صفر والحد الأقصى هو العدد المساوي لعدد فقرات الاختبار. وتعد طريقة إعطاء درجات أو أوزان مختلفة للأسئلة الأكثر أو الأقل صعوبة أو لخيارات مختلفة للسؤال ذاته طريقةً منطقية وعادلة إلا أنه وفي الاختبارات التي تتجاوز فقراتها العشرين لا يكون تأثير مثل هذه الاختلافات في الأوزان محسوساً على تمييز الفقرات الاختبارية أو على ثبات الاختبار. كما لم يتم إثبات أن ما يسمى بوزن الثقة Confidence weighting (حيث توزع الدرجات على أساس نقاط الثقة التي يعطيها الممتحن لإجابته والتي تعبر عن درجة ثقته في صحة إجابته) هو أكثر دقة من المقاييس التقليدية التي تعتمد على مبدأ صفر - ١ في وضع الدرجات. ومن أكثر طرق وضع الأوزان شيوعاً الطريقة المستخدمة في نظرية الاستجابة للأسئلة والتي تعتمد على النظر إلى درجات الاختبارات على أنها تعبير عن درجات تقديرية تقع ضمن مجال القدرة الكامنة. كما تقوم بعض الاختبارات مثل اختبارات الذكاء الفردية (كمقياس ويكسلر لذكاء البالغين) بإعطاء درجات مختلفة (٠ أو ١ أو ٢) للإجابات وفقاً لدقتها أو شموليتها أو لسرعة التوصل إليها.

ويرتبط إجراء "التصحيح من أثر التخمين" باختلاف أوزان الفقرات المختلفة. ويطبق هذا التصحيح على عدد الإجابات الصحيحة في الاختبارات الموضوعية باستخدام المعادلة التالية:

$$s = r - w/(k-1) \qquad (١-٣)$$

وتستند هذه المعادلة إلى الاعتقاد البسيط بأن الممتحَنين الذين لا يعرفون الإجابة على سؤال ما سوف يقومون باختيار واحد من الخيارات عشوائياً. وفي المعادلة تعبر s عن درجات الاختبار بعد تصحيح التخمين، و r عدد الأسئلة التي أجاب عليها الممتحن بشكل صحيح، و w عدد الإجابات الخاطئة، و k عدد البدائل لكل سؤال. وقد تم اقتراح معادلات أخرى لتصحيح التخمين بسبب الاعتقاد بأنه لا يجب أن يلعب التخمين العشوائي دوراً في رفع درجات الأفراد. فعلى سبيل المثال تبدو معادلة ليتل (١٩٦٢) على النحو التالي: s =r-w/[2(k-1)]، ومعادلة ريد (١٩٧٧) s = (n+r) (kr - n)/[2r(k-1)]، حيث n هي العدد الإجمالي لفقرات الاختبار. لا تتضمن معادلة ريد عدد الإجابات الخطأ ويمكن أن ينتج عنها معدل درجات أعلى من ذلك الناتج من المعادلة التقليدية التي فيها s = r. ويساعد برنامج B-3 الموجود على القرص المرافق لهذا الكتاب في حساب التصحيح من أثر التخمين وفقاً للمعادلة المتعارف عليها ومعادلات ريد وليتل.

وتم استكشاف طريقة أخرى للأسئلة الموضوعية وهي إجراء وضع الدرجات عن طريق الاستبعاد وينطوي التغيير في هذا الأسلوب على تعيين درجتين؛ واحدة تحدد للخيارات الصحيحة وأخرى للخيارات الخاطئة، ثم تحسب الدرجة النهائية باعتبار وزنين متساويين من الدرجتين (إدوارد وهاريس، ١٩٨٣). وعلى الرغم من أن كوليت (١٩٧١) وجد أن طريقة وضع الدرجات عن طريق الاستبعاد هي أكثر

صدق وثبات من طريقة التصحيح من أثر التخمين أو الأوزان المختلفة لفقرات الاختبار، وجد هاكستيان وكانسوب (١٩٧٥) أن اختلاف الأوزان المعطاة للإجابة على الفقرات الاختبارية لا يزيد من صدق أو ثبات الاختبار.

أسئلة إعادة الترتيب

توضع درجات اختبارات المطابقة عن طريق إعطاء درجة واحدة لكل مطابقة صحيحة ودرجة الصفر للمطابقات الخاطئة. أما في حالة فقرات الاختبار التي تتطلب إعادة ترتيب محتويات السؤال أو تدريجها تسلسلياً فيصعب استخدام هذه الطريقة في وضع الدرجات التي تعتمد على مبدأ كل شيء أو لا شيء (أي أن الإجابة إما خاطئة تماماً أو صحيحة تماماً)، إذ على الرغم من أن عدد العناصرالمطابقة والمراد ترتيبها يمكن أن يكون نفسه في سؤالين مختلفين من الممكن أن يكون عدد العناصر التي وضعها الممتحَن في المكان الخطأ أقرب إلى عدد العناصر كلها. ولهذا السبب ينصح باستخدام المعادلة التالية عندما يطلب من الممتحَنين ترتيب مجموعة من العناصر (الأشياء أو الأحداث) m بناءً على صفة معينة (كالحجم أو العمر أو التاريخ... إلخ):

$$s = m[1 - 2\sum \mid d \mid /(m^2 - j)] \qquad (٢-٣)$$

وفي هذه المعادلة m هو عدد الأشياء التي تم إعادة ترتيبها وd هو الفرق بين الترتيب الصحيح والترتيب الذي اختاره الممتحَن و0 = j إذا كان m عدداً زوجياً و= ١ j إذا كان m عدداً فردياً و s هو معدل درجات الممتحَن للفقرة الاختبارية. يساعد البرنامجان B-4 أو B-5 على استخدام المعادلة ٣-٢ ومعادلة بديلة (راجع آيكين، a١٩٩٧، ص ٥٩)، للحسابات التي تعطي وزناً أكبر للفروقات الأكبر في المراتب (أي تقتطع درجات أكثر من مجموع الممتحن عندما يرتب العناصر المراد ترتيبها بشكل بعيد جداً عن الترتيب الصحيح).

الاختبارات المقالية

تعد الاختبارات المقالية اختبارات صعبة التصحيح وسهلة التصميم على عكس الاختبارات الموضوعية التي يسهل تصحيحها ويصعب تصميمها. وقد جرى مؤخراً العمل على تطوير طريقة لتقييم الإجابات المقالية عن طريق إجراءات كمّية تتعرف على الملامح الرئيسة للإجابات إلا أن هذه الطريقة لا تزال مكلفة وغير موثوقة ولا تسمح، في الوقت الراهن على الأقل، بالاستغناء عن المصحح التقليدي.

تكمن صعوبة تقييم الأسئلة المقالية في كون هذه الأسئلة ذاتية إذ لا يفهم جميع الممتحنين السؤال بالشكل نفسه كما أن لكل ممتحَن أساليبه المختلفة في الكتابة والتعبير عن الأفكار. ولهذا السبب تختلف محتويات الإجابات وأساليب الكتابة إختلافاً شديداً من شخص إلى آخر. حتى عندما تتوجه الجهود لجعل وضع درجات للاختبارات المقالية أكثر موضوعية وذلك بتحديد أوزان معينة لعناصر المقال المختلفة، فإن هذه العملية تظل معقدة ورتيبة وتتطلب وقتاً كبيراً وليست على درجة عالية من الصدق. مع هذا فإن الدرجات التي تستنبط من مثل هذه العملية التحليلية لوضع الدرجات تعد طريقة أكثر ثباتاً من الطريقة الشمولية في وضع الدرجات والتي تنظر إلى المقال بشكل عام بدلاً من النظر إلى عناصره المختلفة. يضع برنامج ستانفورد لتقييم الكتابة (الطبعة الثالثة) درجات لمقالات الطلاب على مقياس مؤلف من ست نقاط بناءً على نظرة شمولية للمقال برمته. كما يحتوي الكتاب أيضاً على نظام وضع الدرجات التحليلي الذي يعتمد على تقسيم ما يُقيِّم المقال على أساسه إلى ست مجموعات هي: الأفكار والتوسع فيها؛ التنظيم والإنسيابية والتناغم؛ اختيار المفردات؛ الجمل والمقاطع؛ النحو والدقة اللغوية؛ وتقنيات المقال.

ويمكن تدريب المصححين أو قارئي الإجابات (كما يسميهم برنامج المجلس الأعلى للقبول الجامعي) لكي يصبحوا أكثر موضوعية وأكثر ثباتاً على مبادئ وضع الدرجات. وقد درجت الجامعات والكليات المختلفة منذ زمنٍ بعيد وقبل أن تظهر الاختبارات ذات أسئلة الخيارات المتعددة على استخدام الأسئلة المقالية لتحديد الطلاب الذين سيتم قبولهم في هذه الجامعة أو الكلية. ولا تزال خدمة الاختبارات التعليمية Educational Testing Services تستعين بعدد كبير من المدرسين والخبراء سنوياً لتقدير الإجابات على الأسئلة المقالية لاختبارات القبول الجامعي المتقدمة التي يخضع لها طلاب المدارس الثانوية في شتى المواد وفي مختلف أنحاء الولايات المتحدة. يجتمع هؤلاء الخبراء كل عام في مدينة برينستون في ولاية نيوجرسي حيث يتم تدريبهم على ما يتعين عليهم البحث عنه في إجابات الطلاب على اختلاف مستويات دقتها، ويشرف على هؤلاء الأفراد "رئيس المصحّحين" الذي يعطي تعليماته بين الحين والآخر لإعادة قياس طريقة التصحيح وإعادة النظر في المقاييس المعتمدة. ويمكن اتباع إجراء مماثل على مستوى المدرسة وبين مدرسي الفصول المختلفين إذ يتم اختيار إجابات للأسئلة المقالية تنتمي إلى عدد من مستويات التحصيل قبل البدء بعملية وضع الدرجات. ويمكن في حال استخدام طريقة وضع الدرجات الشمولية إعطاء درجة عامة للمقال بين ١ و٥ (كما في اختبارات القبول الجامعي المتقدمة المذكورة أعلاه). أما في حال استخدام طريقة وضع الدرجات التحليلية، فإن تقديرات منفصلة تعطى لعناصر مختلفة من الإجابة المراد تقييمها. وفيما يلي مقترحات أخرى يمكن أن ترفع مستوى صدق وثبات الاختبارات المقالية:

١- يجب وضع درجات سؤال ما على أوراق جميع الممتحَنين قبل الإنتقال إلى السؤال التالي.

٢- يجب وضع درجات السؤال ذاته على أوراق جميع الممتحنين خلال المدة الزمنية ذاتها.

٣- يجب تقييم الناحية الشكلية (كالأسلوب وتقنيات الكتابة ونوعيتها) بمعزل عن تقييم المحتوى في الحالات التي يتم تقييم المقال فيها على أساس هذين العاملين.

٤- يزداد ثبات الدرجات عندما يتم حسابها كمعدلات وسطية لدرجات وضعها مصححان أو أكثر.

٥- يجب تصحيح الأخطاء على ورقة الإجابة وكتابة تعليق إلى جانب كل إجابة.

تقييم الدرجات وتفسيرها

لا تعني درجات الاختبارات أي شيء إن لم تتم مقارنتها مع توقعات أو مقاييس أو معايير معينة. ففي حالة الاختبارات ذات المرجعية المحكية تتم مقارنة درجات الأفراد مع المعيار المتفق عليه والذي يحدد مستوى الأداء الفعال ثم يتم تحويل الدرجات الأولية إلى نسب مئوية تحدد موقع الطالب بالنسبة إلى المعيار العام. ويتم عادةً تحديد المعيار العام لاختبار تحصيلي ذو مرجعية محكية من قبل لجنة مكونة من عدد من المدرسين أو الإداريين في المدارس أو خبراء التقييم أو غيرهم من الأفراد المطلعين على التقدم في التعليم وأساليب تقييمه.

وعلى عكس الاختبارات ذات المرجعية المحكية تُقارن معدلات الدرجات في الاختبارات ذات المرجعية المعيارية بعدد من المعايير التي تم التوصل إليها عن طريق دراسة التوزيع التكراري لدرجات تم الحصول عليها من عينة تمثل مجموع الأفراد الذين طور الاختبار من أجلهم. وكما ذكرنا في الفصل الثاني فإن المعدلات أو الدرجات المحوَّلة يمكن أن تكون عمراً أو مكافئات للدرجة أو مراتب مئينية أو درجات قياسية. ومن الجدير بالذكر هنا أنه، وعلى عكس مقاييس الاختبارات ذات المرجعية المحكية، فإن المعايير ليست مقاييس باعتبار كونها أهدافاً أو متطلبات راسخة؛ بل مجرد معدلات وسطية. وتلخيصاً لما سبق يمكن

القول أنه يتم مقارنة درجات اختبار ذو مرجعية محكية بمستوى تم الاتفاق عليه واعتباره مستوى مرغوب فيه للأداء بينما تتم مقارنة درجات اختبار ذو مرجعية معيارية بأداء مجموعة من الأفراد مشابهة للممتحَنين بطرق واضحة.

ولا تعتبر مقاييس الأداء الذاتية التي يضعها المدرسون والمقيمون والمرشدون الاجتماعيون مقاييس ثابتة منقوشة على الحجر بل هي مجرد ملاحظات خلص إليها هؤلاء الأشخاص نتيجة إحتكاكهم بالطلاب. وعليه فإن المقاييس التي يعتمدها المدرسون والأخصائيون الاجتماعيون وأرباب العمل لفهم درجات الأفراد هي مقاييس ضمنية وغير صريحة. وتؤثر عوامل أخرى على هذه المقاييس الذاتية كالتقييم الذاتي وتوقعات المقيِّم أو طموحاته لأداء الممتحنين.

من غير الممكن فهم درجات الاختبارات (في المدارس أو مراكز الإرشاد أو العيادات أو الشركات والمؤسسات) بمعزل عن معلومات متعلقة بخلفية الممتحنَين، إذ يجب الإستعانة بمعلومات عن العمر والمستوى التعليمي والصفات الجسدية والحالة الصحية والخلفية الحضارية والإنجازات السابقة للفرد وغيرها من العوامل المتعلقة بالخبرة والأمور البيولوجية لتفسير درجات الأفراد قبل إتخاذ قرارات تخصهم في مجالات التعليم والتوظيف والعلاج. ومن المهم أن نكون على قناعة تامة بأن نتيجة الاختبار ليست علامة تطبع على جبين الممتحَن لتلازمه طوال حياته. فكل درجة يتم الحصول عليها تحتوي على نسبة من الخطأ التي تختلف مع الوقت، ولذلك تعتبر الدرجات إشارات أو إقتراحات لا يكتمل معناها إلا بالنظر إلى معلومات أخرى عن الممتحَن. وإضافةً إلى ذلك من المفيد مقارنة درجات الأفراد ليس مع الآخرين فحسب بل مع أنفسهم حيث تساعد معلومات عن درجات ممتحَن ما في عدد من الاختبارات الفرعية لسلسلة من اختبارات التحصيل أو الاستعدادات المتعددة على تكوين صورة واضحة عن نقاط ضعف وقوة هذا الممتحن المختلفة في مجال الإدراك.

يصنف البشر الأحداث والأشياء والأفراد وكل ما يشعرون به أو يحسونه بالفطرة ودون أن يجبروا على هذ التصنيف. ويصنف علماء النفس الناس بناءً على درجاتهم في اختبارات القدرة واختبارات الشخصية إلى: أفراد طبيعيين أو معاقين عقلياً أو موهوبين أو مضطربين نفسياً أو مصابين بالعصاب... إلخ. ويصنف العاملون في مجال التعليم الأفراد إلى: طلاب ضعيفين أو بطيئي التعلم أو مصابين بنقص الانتباه وفرط الحركة (ADHD) أو مصابين بمرض صعوبة القراءة والكتابة (ديسلكسيا)... إلخ. كما أن هنالك طلاب ممتازون أو ما يسمى طلاب الدرجة A وطلاب جيدون جداً يحصلون على الدرجة B معظم الوقت وطلاب جيدون يحصلون على الدرجة C، إلا أن هذا لا يعني أن هذه الدرجات موحدة في المقررات المختلفة. وقد تكون درجات الأفراد مستخلصة مباشرةً من الدرجات الأولية للاختبارات ومن الممكن أن يكون المصحح قد أخذ مستوى الأداء العام للفصل وتوقعاته من الطلاب بعين الاعتبار. ويساعد برنامج B-6 الموجود على القرص المرافق للكتاب على إدخال جميع العوامل التي تدخل في حساب الدرجات النهائية من التوزيع التكراري لدرجات الاختبار إلى وسيط مستوى قدرة الفصل التقديري (على مقياس من ١ إلى ١٠٠) إلى الدرجات العظمى والدنيا التي حصل عليها الطلاب إلى الدرجات الأولية لكل ممتحن. وبالرغم من كل هذه التقنيات لا يتم وضع الدرجات النهائية لأي ممتحن دون الرجوع إلى خاصياته الاجتماعية والسياسية والتعليمية.

صياغة نتائج الاختبارات في تقارير رسمية

بالاعتماد على طبيعة الاختبار وعلى كيفية استعمال النتائج، فإن الدرجات سيتم رصدها في دفتر الدرجات الخاص بمدرس المادة، أما في حالة الاختبارات القياسية تُسجل نتائج الاختبارات في سجل تراكمي أو في بطاقة التقرير المدرسي. كما تحفظ درجات المتقدمين إلى مؤسسة تعليمية أو مهنية ما مع

غيرها من أوراق الطلب الرسمية للعودة إليها عند اللزوم. ومن الممكن إرسال نتائج الاختبارات مع غيرها من المعلومات عن شخصٍ ما إلى أفراد مؤهلين يودون الإستفادة من هذه المعلومات.

وعند إحالة قضية طفل أو شخص للدراسة من قبل مدرسة أو مركز إرشاد أو عيادة نفسية، تكون تقارير نتائج الاختبارات وغيرها من إجراءات تقييم قدرات الأفراد تقاريراً مستفيضة. وغالباً ما تتألف هذه التقارير من عدة أقسام تضم معلومات الممتحَن الشخصية (كاسمه وعمره وتاريخ ميلاده وخلفيته التعليمية) واسم القائم على الاختبار ومكان الاختبار وتاريخه ونوع الاختبارات التي أجريت سابقاً ونتائجها بالإضافة إلى الاستنتاجات والتوصيات. ويجب أن تحمل التقارير توقيع المشرف على الاختبار وتاريخ تحرير التقرير. ويجب إعارة قسم الاستنتاجات والتوصيات اهتماماً خاصاً إذ يضم هذا القسم النتيجة النهائية لعمليتي التقييم والتشخيص. ويجب أن تتم صياغة التقريربأسلوب متزن يعبر بوضوح عن الاتجاه النظري للأخصائي الذي درس الحالة والفرضيات النفسية التي بنيت الاستنتاجات النهائية عليها. كما يجب توخي الحذر من حيث التعبير باستخدام لغة الاحتمال بدلاً من لغة التأكيد، وإذا لم يكن بالإمكان الحصول على استنتاجات يمكن تبريرها من خلال النتائج فيجب أن لا تذكر أي استنتاجات.

الخلاصة

تناول هذا الفصل من الكتاب عمليات تصميم ووضع وإجراء الاختبارات ثم تصحيحها وتحليل نتائجها وكتابة تقارير عنها. ويجب قبل البدء بعملية وضع الاختبار تحديد الهدف من الاختبار والمجموعة البشرية المراد اختبارها ووقت الاختبار وطريقة التصحيح وكيفية استخدام نتائج الاختبار. يمكن تصميم اختبارات الاستعداد لتعمل كمقاييس تراكمية أو تكوينية للتحصيل إذ يمكن لهذه

الاختبارات أن تساعد في توزيع الأفراد على الصفوف الدراسية أو الوظائف المختلفة أو أن تساعد في تشخيص ومعالجة بعض المشاكل والاضطرابات أو في تقييم البرامج والمشاريع والمؤسسات. ويتوجب على مستخدم الاختبار التأكد من أن الاختبار مناسب للهدف المراد استخدامه لتحقيقه.

يساعد جدول مواصفات الاختبار أو المخطط التفصيلي للاختبار على تحديد شكل ونوع ومحتوى الاختبار. وعلى الرغم من أن فقرات الاختبار الموضوعية كأسئلة الخيارات المتعددة وأسئلة الصواب والخطأ هي الأسئلة الأكثر شيوعاً في الاختبارات القياسية، تعتبر الاختبارات ذات الأسئلة المقالية مقاييس أفضل للعديد من الأهداف. وإن ترتيب فقرات الاختبار عن طريق وضع الأسئلة المتشابهة في القسم ذاته وترتيب الأسئلة من الأسهل إلى الأصعب قد يكون أمراً مناسباً للممتحَنين والقائمين على الاختبارات، إلا أن مصممي الاختبارات سيحسنون صنعاً عندما يركزون أكثر على كتابة فقرات اختبار جيدة أكثر من تركيزهم على ترتيبها. ويجب أن تكون تعليمات الاختبارات، التي يمكن أن يقرأها القائم على الاختبار شفهياً أو أن يقرأها الممتحَنون، واضحة وكاملة.

تتنوع المؤهلات المطلوبة من المسؤولين على الاختبارات وتكون المتطلبات في حالات الاختبارات الفردية واختبارات الذكاء والشخصية أكثر صرامة من متطلبات الاختبارات الجماعية. ويتعين على دور نشر وتوزيع الاختبارات التأكد قدر الإمكان من توفر الحد الأدنى من مؤهلات إجراء الاختبارات لدى الأفراد الراغبين في شراء الاختبارات منهم. ويُطلب من القائمين على الاختبارات أن يكونوا مسؤولين ليس عن الذي يجري أثناء فترة الاختبار فحسب بل عن التحضير للاختبار من حيث الناحية الأمنية والظروف المواتية قبل بدء الاختبار وعن التعامل الصحيح مع نتائج الاختبار بعد انتهائه. قد يحدث الغش في أي مرحلة من مراحل الاختبار (أي قبله أو أثناءه أو بعده) ولا يجب الخلط بين الغش

وبين البراعة في الاختبارات التي تعني أن بعض الأفراد موهوبون من حيث قدرتهم على تحقيق مستويات أداء عالية في جميع الاختبارات.

يجب الالتزام بطريقة إجراء الاختبار المحددة وخصوصاً في حالة الاختبارات القياسية. إلا أنه هنالك حالات تتطلب عدم التقيد بهذه الإجراءات مثل حالات اختبار الأفراد المعاقين الذين لا يمكنهم الالتزام بتعليمات الاختبار العادية. ولكن من المهم هنا مراعاة عدم مخالفة تعليمات الاختبارات إلا إذا كان الهدف الحصول على معلومات أكثر دقة فيما يخص بعض الممتحَنين، كما يجب مراعاة عدم مقارنة نتائج الحالات الإستثنائية مع المعايير المستخلصة من عينات المعايرة الطبيعية.

ويتم إجراء معظم الاختبارات الفردية بالطريقة التقليدية التي تتطلب ممتحن وواحد وشخص قائم على الاختبار، بينما تستخدم الاختبارات الجماعية الورقة وقلم الرصاص. وقد ازدادت مؤخراً شعبية الاختبارات التي تستعين بالحاسب وخاصة الاختبارات التكيفية (CAT) مثل اختبار التقييم المدرسي (SAT) واختبار سجل الخريجين (GRE) وسلسلة اختبارات القوات المسلحة للقابلية المهنية والتي تُجرى للآلاف من الأفراد داخل وخارج الولايات المتحدة.

وتعد الطريقة الأكثر استخداماً في وضع درجات الاختبارات الجماعية الموضوعية كالاختبارات التحصيلية واختبارات الاستعداد طريقة إعطاء درجة واحدة للإجابة الصحيحة ودرجة الصفر للإجابة الخاطئة. أما في اختبارات إعادة الترتيب والاختبارات المقالية فتعطى درجات مختلفة للإجابات بناءاً على دقتها أو شموليتها. وعلى الرغم من أن التصحيح من أثر التخمين لا يزال إجراء يرتبط بالعديد من الاختبارات، وكقاعدة عامة لا يعتبر استخدام مثل معادلات تصحيح التخمين هذه مع الاختبارات المدرسية خياراً صائباً.

يتم التعبيرعن درجات الاختبارات ذات المرجعية المحكية على شكل نسب مئوية بالنسبة إلى محك القبول أو مستوى الإتقان. أما في الاختبارات ذات المرجعية المعيارية فيتم تحويل الدرجات إلى معايير للعمر أو الصف أو معايير مئينية أو درجات معيارية تم الحصول عليها من نتائج عينة من الأفراد تمثل المجموعة البشرية المراد اختبارها. ويتوقع أن تكون عملية التعبير عن معدلات درجات الاختبارات عملية موضوعية إلا أنها في الواقع عملية ذاتية ترتبط بخبرة وتوقعات والتوجهات النظرية للشخص الذي يقوم بهذه العملية. وغالباً ما يتم تحويل درجات الاختبار في المدارس إلى درجات واستخدامها في تقرير نجاح أو رسوب الطالب.

يتم تدوين نتائج الاختبارات المدرسية في سجلات تراكمية أو في التقارير المدرسية. ويحدث شيء مشابه في مؤسسات التعليم العالي وبعض المؤسسات المهنية. أما في حالة الاختبارات النفسية والتشخيصية فيجب إعداد تقرير تفصيلي يضم معلومات عن شخصية الفرد الذي تم اختباره والسبب الذي أدى للاختبار ونتائج المقابلات والملاحظات والنتائج المحددة للاختبار النفسي بحد ذاته. ويختتم التقرير عادةً بقائمة من الاستنتاجات والتوصيات وذلك لغرض التقدير التالية أو طرق المعالجة أو غيرها من الإجراءات.

obeikandl.com

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

الاختبارات التحصيلية

وضع الدرجات التحليلي

اختبارات الاستعداد

تقنية حذف الكلمات من النصوص (cloze)

الاختبار التكيفي المحوسب (CAT)

الوزن المرتبط بالثقة

وضع الدرجات عن طريق الاستبعاد

أسئلة إعادة الترتيب

أسئلة الإجابات القصيرة

أدوات الدلالة المحددة

الكتيب الامتحاني اللولبي

السؤال المقالي

الاختبار ذو المستويات المرنة

وضع الدرجات الشمولي

أسئلة المطابقة

الاختبار متعدد المستويات

أسئلة الخيارات المتعددة

الاختبار الخارج عن المستوى

التواصل

تصنيف الأهداف التعليمية

البراعة في الاختبارات

أسئلة الصواب والخطأ

obeikandl.com

المراجع المقترحة

Aiken, L. R. (1991). Detecting, understanding, and controlling for cheating on tests. *Research in Higher Education, 32,* 725–736.

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.

Golden, C. J., Sawicki, R. F., & Franzen, M. D. (1990). Test construction. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 21–40). New York: Pergamon Press.

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.

Gronlund, N. E. (1993). *How to make achievement tests and assessments* (5th ed.). Boston: Allyn & Bacon.

Haladyna, T. M. (1994). *Developing and validating multiple-choice items.* Hillsdale, NJ: Erlbaum.

Lehman, R. (1994). Essays, scoring of. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 4, pp. 2018–2024). Tarrytown, NY: Elsevier.

Millman, J., & Greene, J. (1989). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 335–366). New York: American Council on Education/Macmillan.

Nitko, A. J. (1989). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 447–474). New York: Macmillan.

Robertson, G. J. (1990). A practical model for test development. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence & achievement* (pp. 62–85). New York: Guilford Press.