جامعة الإمام عبدالرحمن بن فيصل
IMAM ABDULRAHMAN BIN FAISAL UNIVERSITY
وكالة الجامعة للشؤون الأكاديمية
Vice Presidency for Academic Affairs

مركز جودة التقويم والاختبارات
Center For Quality of Assessment and Examinations

دليل إعداد الاختبارات التحصيلية

الإصدار الأول ٢٠٢٢-٢٠٢٣

جدول المحتويات

مقدمة..... ٤

تعريفات..... ٥

أغراض الاختبارات التحصيلية..... ٦

دور الاختبارات التحصيلية في عملية التعليم والتعلم..... ٦

أسس تصنيف الاختبارات..... ٧

خصائص الاختبار الجيد (الخصائص السيكومترية للاختبارات)..... ٩

صدق الاختبارValidity..... ٩

طرق التحقق من صدق الاختبار التحصيلي..... ٩

ثبات الاختبار التحصيلي Reliability..... ١٢

طرق حساب ثبات الاختبار..... ١٣

العلاقة بين الثبات و الصدق..... ١٦

الموضوعية Objectivity..... ١٧

التوازن Balance..... ١٧

العدالة Fairness..... ١٧

صلاحية الاستخدام Usability..... ١٧

خطوات إعداد الاختبار التحصيلي الجيد..... ١٨

أولًا: تحديد الهدف من الاختبار..... ١٨

ثانيًا: الموائمة مع مخرجات التعلم المراد قياسها..... ١٩

ثالثًا: تحليل المحتوى..... ٢٢

رابعًا: بناء جدول المواصفات..... ٢٣

خامسًا : صياغة المفردات الاختبارية..... ٢٦

سادسًا: إخراج الاختبار..... ٢٧

سابعًا: تطبيق الاختبار..... ٢٩

ثامنًا: تصحيح الاختبار..... ٢٩

تاسعًا: تحليل نتائج الاختبار..... ٣٠
عاشرًا: استثمار نتائج الاختبار..... ٣٠
الاختبارات الشفوية..... ٣١
أسئلة شائعة..... ٣٤
المراجع..... ٣٧

مقدمة

يهدف هذا الدليل إلى دعم أعضاء هيئة التدريس، ومن في حكمهم في بناء الاختبارات التحريرية والشفوية بهدف قياس مخرجات التعلم الخاصة بمجالات المعرفة والفهم، والمهارات الإدراكية، ومهارات التواصل وفقًا لمعايير بناء أدوات التقويم، وبما يحقق سياسات وإجراءات جودة التقويم والاختبارات، وتجويد ممارسات تقويم الطلاب والاختبارات لدى أعضاء هيئة التدريس بكليات جامعة الإمام عبد الرحمن بن فيصل كأحد الأهداف الاستراتيجية لمركز جودة التقويم والاختبارات.

يتناول هذا الدليل بالشرح والتوضيح مراحل بناء الاختبارات التحريرية، وطرق التحقق من خصائصها السيكومترية مثل الثبات والصدق، وغيرها من الخصائص التي تضمن تقديم أدلة على سلامة القرارات المستندة إلى نتائج أدوات التقويم المستخدمة بالبرامج الأكاديمية، وبخاصة الاختبارات، وضمان كفاءة عمليات تقييم الطلاب بجامعة الإمام عبد الرحمن بن فيصل.

تعريفات

الاختبار Exam

إجراء منظم أو أداة قياس مقننة للحصول على قياس موضوعي لعينة من السلوك، بهدف مقارنة أداء الفرد بمعيار أو بمستوى أداء محدد لإظهار عينة من سلوك الفرد.

الاختبار التحريري Written Exam

إجراء منظم للحصول على قياس موضوعي لعينة من إنجاز الطالب لمخرجات التعلم في مجالات المعرفة والفهم، والمهارات المعرفية، وفقًا للإطار الوطني للمؤهلات، ويتم بناؤه وفقًا لجدول المواصفات، بحيث تواءم مفرداته الاختبارية مخرجات التعلم وأوزانها النسبية.

المفردة الاختبارية Test item

وحدة بناء الاختبار، وهي أصغر وحدة يتألف منها الاختبار، على سبيل المثال: قد يشتمل الاختبار على سؤال من نمط الاختيار من متعدد يتضمن ٣٠ مفردة اختبارية.

صدق الاختبار Validity

أن يقيس الاختبار ما وضع لقياسه، وأن يحقق هدف القياس الذي وضع لأجله.

صدق المحتوى Content Validity

تمثيل مفردات الاختبار للمجال الذي وضع لقياسه (محتوى المقرر ومخرجات التعلم).

الثبات (الموثوقية) Reliability

اتساق نتائج الاختبار واستقرارها، وتجنب أخطاء القياس، مما يؤدي إلى الحصول على نفس النتائج عبر تطبيقات متعددة، ويعبر عنها باحتفاظ الأفراد بالمركز النسبي لهم (ترتيبهم) في المجموعة.

المحك Criterion

درجة تُستخدم في تفسير أداء الطالب على الاختبار، وتُحدد سلفًا في ضوء هدف أو مستوى ينبغي الوصول إليه، مثل تحديد نسبة ٦٠٪ كمتطلب للنجاح.

المعيار Norm

درجة تُستخدم في تفسير أداء الطالب على الاختبار، وتُحدد في ضوء أداء المجموعة بعد تطبيق أداة التقويم وتحليل النتائج. مثل: المتوسط، الوسيط، المنوال (مقاييس النزعة المركزية)، أو المئينيات(Percentiles)

التعلم Learning

تغير شبه دائم في سلوك الفرد نتيجة اكتساب مجموعة من المعارف والمهارات والاتجاهات يُعبر عنها إجرائيًا باكتساب وإنجاز مخرجات التعلم.

مخرجات التعلم Learning Outcomes

تعبر عن السلوك المتوقع من الطالب أدائه بنهاية عملية التعليم والتعلم في المقرر، ويتم تصنيفها في توصيف المقرر في مجالات ثلاثة هي المعرفة والفهم، والمهارات، والقيم، وذلك طبقاً للإطار الوطني للمؤهلات ٢٠٢٠.

الإنجاز الأكاديمي Academic Achievement

النجاح في أحد مجالات الدراسة ويعبر عنه بمستوى من الكفاءة في العمل الدراسي بشكل عام أو في مهارة معينة، والذي يستدل عليه من نتائج الاختبارات المقننة، وأداء الطلاب في أدوات التقييم من إعداد المعلم (أستاذ المقرر).

قلق الاختبار Test Anxiety

التوتر والقلق المرتبطان بإجراء الاختبار، مما يؤدي في كثير من الأحيان إلى انخفاض أداء الطالب في الاختبار.

دافعية التعلم Learning Motivation

استعداد الشخص لبذل جهد بدني أو عقلي سعياً لتحقيق أهداف التعلم.

أغراض الاختبارات التحصيلية

- **التشخيص:** التعرف على جوانب القوة والضعف في أداء الطالب، وذلك بهدف تعزيز جوانب القوة، ومعالجة جوانب الضعف.
- **التصنيف:** تصنيف المتعلمين إلى مستويات أو إلى تخصصات من خلال الكشف عن الفروق الفردية في أدائهم الأكاديمي.
- **قياس نسبة الإنجاز:** تحديد نسبة تحقيق الطالب لمخرجات تعلم المقرر، وترفيعه إلى صفوف أعلى، وتحديد نسبة إنجاز أستاذ المقرر لمخرجات التعلم، وجودة طرق تدريسه، واستراتيجياته، وممارساته، داخل قاعات الصف الدراسي.
- **التنبؤ:** تحديد مستوى أداء الطالب في المراحل المستقبلية من التعلم، أو تحديد مستوى تنافسيته في سوق العمل بعد التخرج.

دور الاختبارات التحصيلية في عملية التعليم والتعلم

تمثل الاختبارات التحصيلية أحد أدوات التقويم بأنواعه المختلفة: التشخيصي، البنائي، الختامي، وعليه تسهم الاختبارات في تعزيز عملية التعليم والتعلم من خلال التالي:

- **الاختبارات كأداة قياس لمستوى المتعلم Assessment of Learning**

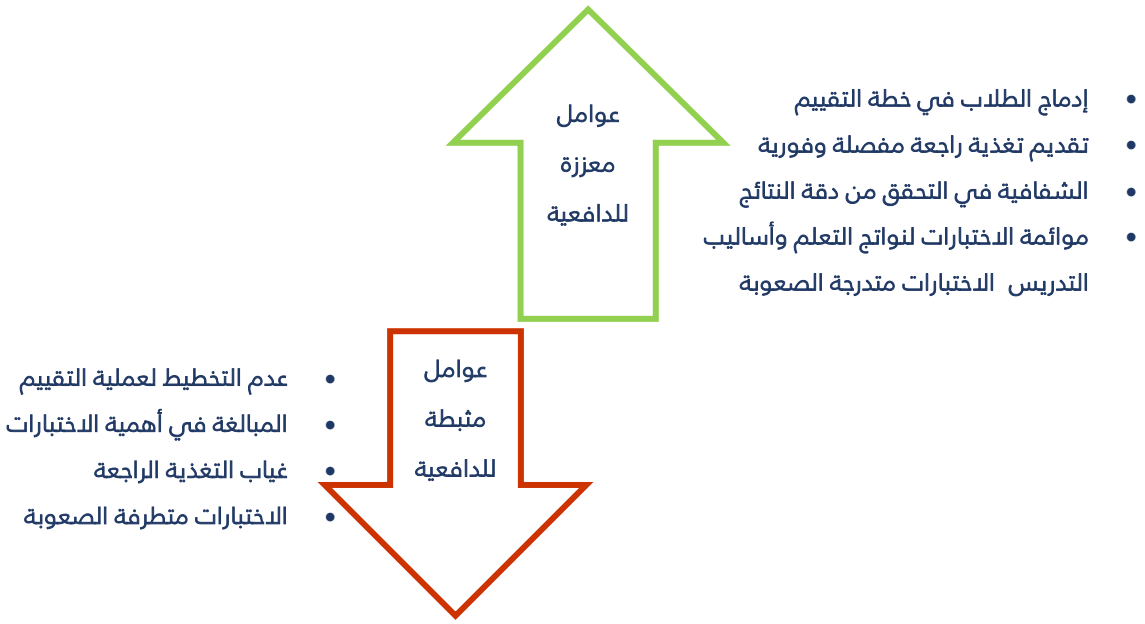
تمثل الاختبارات أحد أدوات التقويم الختامي Summative Assessment، حيث تقيس مستوى تمكن الطلاب من تحقيق مخرجات التعلم من خلال معايير محددة، والكشف عن الفروق الفردية بينهم، واتخاذ القرارات المستندة إلى ذلك مثل: استخدام درجات الاختبارات لأغراض ترفيع الطلاب إلى مستويات أعلى، أو انتقاء المتفوقين، أو تحديد المتعثرين لتقديم الدعم الأكاديمي المناسب لهم، ويمثل تحقيق الطلاب لمخرجات التعلم مؤشرًا على جودة العملية التعليمية.

❖ الاختبارات كأداة لتحسين وتعزيز التعلم Assessment for Learning/Assessment as Learning

تستخدم الاختبارات كأحد أدوات التقويم البنائي أو التكويني، حيث تعزز التعلم، وذلك من خلال إمداد الطلاب بتغذية راجعة بنائية ومفصلة عن أدائهم بما يمكن الطالب من متابعة تقدم أدائه وتحسينه، بالإضافة إلى تقييم أستاذ المقرر لطرق واستراتيجيات التدريس التي يتبناها، والحكم على فاعليتها في تحقيق مخرجات التعلم المستهدفة، وتعزيز عملية التعليم والتعلم.

❖ الاختبارات كأداة لزيادة دافعية المتعلمين

تؤثر الاختبارات التحصيلية في زيادة دافعية الطلاب من خلال دمجهم ومشاركتهم في عملية التقييم؛ وذلك عن طريق إعلامهم بخطة التقييم، وطبيعة مهام التقييم مع بداية الفصل الدراسي، والمشاركة في بناء مصفوفة المعايير/ التصحيح، وتحديد معايير التقييم، واستطلاع آرائهم في مواعيد الاختبارات، مع التأكيد على استيفاء أسس المواءمة البنائية التي تضمن المشاركة الفاعلة للطالب، وتعزيز أسس التعلم الذاتي مما يُشعِر الطالب بأنه متحكم بعوامل نجاحه، وتتولد لديه مشاعر إيجابية مرتبطة بالاختبارات مثل الشعور بالإنجاز؛ مما يزيد من دافعيته للتعلم. ومن ناحية أخرى عندما يُضاف إلى خطة التقييم مهام لم يسبق تضمينها بشكل مفاجئ للطلاب، أو التطرف في صعوبة المفردات الاختبارية، أو أن تقيس الاختبارات مخرجات تعلم لم يتم تغطيتها خلال مدة التدريس، أو أن يقوم أستاذ المقرر بالمبالغة في أهمية الاختبار مما يزيد من مشاعر القلق والإحباط لدى الطلاب بما يؤثر سلباً في دافعيتهم للتعلم. ويوضح الشكل التالي علاقة ممارسات التقييم ودافعية الطلاب.



شكل١. علاقة ممارسات التقييم ودافعية الطلاب

أسس تصنيف الاختبارات

تتعدد أنواع الاختبارات التحصيلية وتتنوع استخداماتها في العملية التعليمية، ويفيد تصنيف الاختبارات في مساعدة عضو هيئة التدريس على معرفة نوع الاختبار الذي يقدمه، ونطاقه، ومحدوديته، ومميزاته.

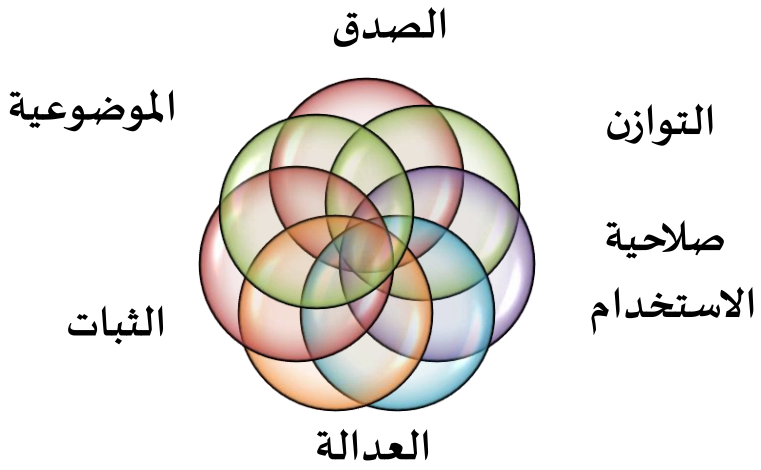
الأساس التصنيفي		التعريف
الجهة المعدة للاختبار	مقننة	اختبارات معدة من قبل متخصصين/ هيئات أو مراكز القياس والتقويم المعتمدة، ويتم فيها وضع معايير للدرجات يُرد إليها درجات الأفراد، مثل: الاختبارات المهنية واختبارات اجتياز اللغات الأجنبية: مثل اختبارات TOEFL واختبارات الهيئة الوطنية السعودية للتخصصات الطبية.
	إعداد الأستاذ	اختبارات يعدها أستاذ المقرر، ويطبقها، ويصححها، ويحلل نتائجها، وقد يتحرر فيها من بعض الخطوات والعلميات مثل التجريب أو التقنين، إلا أنها تتضمن التحقق من الخصائص السيكومترية مثل الصدق من خلال تحليل محتوى المقرر، وإعداد جدول المواصفات ومراجعة مدى موائمة الاختبار معه.
تفسير النتائج	معياري المرجع	مقارنة أداء الفرد بمعيار الجماعة التي ينتمي إليها، ويتمثل هذا المعيار بدرجة تُعبر عن أداء المجموعة: مثل: (متوسط- وسيط- منوال)، فعلى سبيل المثال، مقارنة علامة الطالب بمتوسط أداء أقرانه في الشعبة الدراسية أو المقرر الدراسي.
	محكي المرجع	مقارنة أداء الفرد بمحك، يُمثل مستوى أداء محدد سلفًا بغض النظر عن أداء أقرانه أو مجموعته، ويتم تحديد هذا المحك في ضوء هدف أو مستوى ينبغي الوصول إليه، مثل: درجة النجاح ٦٠٪.
طريقة الإجابة	موضوعية	الاختبارات التي تكون فيها الإجابة الصحيحة متوفرة من بين مجموعة من البدائل، حيث يتعرف الطالب على الإجابة الصحيحة. مثل: مفردات الاختيار من متعدد، أو المزاوجة، أو الصواب والخطأ.
	مقالية	الاختبارات التي يطلب من الطالب فيها أن يصوغ أو ينتج إجابة بنفسه، وقد تكون الإجابة مقيدة مثل: الإكمال والإجابات القصيرة أو حرة مثل: الأسئلة المقالية.
عدد الممتحنين	فردية	الاختبارات التي يؤديها كل فرد على حدة مثل الاختبارات الشفوية، واختبارات الأداء العملي، وهنا يكون عدد مرات تطبيقها يساوي عدد الطلاب.
	جماعية	الاختبارات التي يؤديها عدد كبير من الطلاب في نفس الوقت، مثل: الاختبارات الفصلية أو النهائية.

ويمكن تصنيف الاختبارات التحصيلية في ضوء الأسس السابقة إلى التالي:

- من إعداد الأساتذة، فيما عدا الحالات التي تُستخدم فيها الاختبارات المهنية المقننة لتقييم أداء الطلاب في بعض التخصصات.
- محكية المرجع، حيث تُفسر درجات الطلاب في ضوء محكات محددة مسبقًا، مثل: تحديد درجة للنجاح والتي تتمثل في نسبة محددة من الدرجة الكلية (٦٠٪ من الدرجة الكلية للمقرر) أو التقديرات، والتي تعبر عن نسبة تحقيق وإنجاز الطلاب لمخرجات التعلم.
- تمثل الاختبارات التحريرية مثالًا على الاختبارات الجماعية حيث تطبق على شعبة أو مجموعة من الشعب في وقت واحد، بينما تمثل الاختبارات الشفوية مثالًا على الاختبارات الفردية، حيث يُختبر كل طالب على حدة.
- الاختبارات التحريرية بالجامعة هي مزيج بين الاختبارات الموضوعية، والمقالية، حيث تتطلب سياسات وإجراءات التقويم التنوع في طرق التقويم وأدواته، والتنوع في أنماط الأسئلة بما يتواءم مع مخرجات التعلم المراد قياسها.

خصائص الاختبار الجيد (الخصائص السيكومترية للاختبارات)

للتحقق من صلاحية الاختبار التحريري كأداة لقياس ما أنجزه الطالب من مخرجات التعلم المعرفية، ينبغي أن يتم فحص سلامة الاختبار، وخصائصه، المتمثلة بالشكل التالي:



شكل ٢. التأثيرات المتبادلة بين خصائص الاختبار الجيد

صدق الاختبار Validity

يعني صدق الاختبار أن يقيس الاختبار ما وضع لقياسه، أو أن يحقق الغرض الذي أعد من أجله، وإلى أي مدى يمكن الاعتماد على درجاته في إصدار قرارات تربوية أو مؤسسية مهمة، إذ تعتبر عملية القياس صادقة بقدر ما تقدمه من معلومات موثوق بها؛ فيما يتعلق بتحقيق الطالب لمخرجات التعلم. ويعتبر إعداد جدول المواصفات (مخطط الاختبار)، واختيار نوع الأسئلة المناسب خطوة أساسية نحو تحقيق قياس صادق، حيث يرتبط صدق الاختبار بصدق كل مفردة اختبارية فيه، وعلى سبيل المثال: إذا كان مخرج التعلم هو أن يحلل الطالب الأسباب الرئيسة لمرض ما، أو لحدوث تفاعل كيميائي معين، فإن أسلوب القياس الأكثر صدقًا ينبغي أن يتطلب إجابة قصيرة وليس التعرف على الإجابة من خلال أسئلة اختيار من متعدد أو الصواب والخطأ؛ لأن تعرف الطالب على الإجابة، أو تحديد الصواب والخطأ، لا يضمن تحقيقه لمخرج التعلم المستهدف بالقياس، وهو "التحليل".

طرق التحقق من صدق الاختبار التحصيلي

❖ الصدق الظاهري Face Validity

ويُقصد به أن تقيس المفردات الاختبارية ما وضعت لأجله، بحيث تتشابه في محتواها مع المحتوى المستهدف بالقياس، ويتم التحقق الظاهري من خلال عرض أسئلة الاختبار على متخصص في المحتوى الأكاديمي لتحديد درجة

قياس الاختبار للمعارف والمهارات والمحتوى المقاس، ومدى تركيز مفرداته الاختبارية على المحتوى الرئيسي بدلًا من المعلومات المساعدة/ الأقل أهمية، وقد يتم استطلاع رأي الطلاب في مدى ارتباط الاختبار بمخرجات التعلم والمحتوى وذلك بعد أدائه، كأحد طرق التحقق من الصدق الظاهري Face Validity. وعلى الرغم من أن تحقق هذا الجانب من الصدق يعمل على زيادة دافعية الممتحن لأداء الاختبار، إلا أنه لا يكفي للحكم على الاختبار بأنه صادق، ويجب استخدام أدلة إضافية للصدق.

❖ صدق المحتوى Content Validity

ويقصد به كفاية تمثيل عينة الأسئلة المتضمنة في الاختبار للمحتوى، ومدى تركيز هذه الأسئلة على المحتوى الرئيس الذي تتحقق من خلاله مخرجات التعلم المستهدفة بدلًا من المعلومات المساعدة/ الأقل أهمية. وتُعد جودة بناء جدول المواصفات خطوة مهمة لتحقيق صدق المحتوى، **كما يمثل صدق المحتوى الطريقة الأكثر مناسبةً للتحقق من صدق الاختبارات التحصيلية**. ويمكن لأعضاء هيئة التدريس التحقق منه في سهولة ويسر قبل تطبيق الاختبار على الطلاب.

❖ صدق المفهوم (البناء) Construct Validity

يُقصد به **الدرجة التي يكون عندها الاختبار** التحصيلي قادرًا على قياس التحصيل في المقرر المستهدف، فعند إعداد اختبار لقياس التحصيل في مقرر الرياضيات، يجب جمع أدلة على أن الاختبار يقيس التحصيل في الرياضيات فقط، ولا يقيس الذكاء مثلًا، ويُقاس بالعلاقة بين نتائج الاختبارات وبين المفهوم النظري الذي يهدف الاختبار التحصيلي إلى قياسه، ويتم من خلال فحص مفردات الاختبار لمعرفة محتواها باستخدام أساليب إحصائية كالتحليل العاملي على سبيل المثال، أو بحساب معامل الارتباط بين درجات اختبار التحصيل، واختبارات أخري تقيس نفس المجال واختبارات تقيس مجالات مغايرة فيما يعرف بالصدق التمييزي.

❖ الصدق المرتبط بمحك Criterion-related Validity

ويتحقق هذا الصدق من خلال طريقتين:

- ✓ الصدق التلازمي: Concurrent Validity
- ✓ الصدق التنبؤي: Predictive Validity

الصدق التلازمي	الصدق التنبؤي
هو ارتباط الأداء على اختبارين لنفس المفهوم تم تطبيقهما في نفس الوقت أحدهما هو الاختبار التحصيلي والأخر هو اختبار معروف في المجال المُقاس (مثل: الاختبارات المهنية professional exams).	هو ارتباط الأداء على الاختبار التحصيلي بالأداء المستقبلي للطالب مثل: المعدل التراكمي للتقديرات، وأداء الخريج في سوق العمل (رضا جهات التوظيف عن جودة أداء الخريج).

❖ حساب الصدق التلازمي للاختبار

يعبر الصدق التلازمي عن مدى ارتباط نتائج الاختبار بنتائج قياس أخر موثوق ودقيق وذي صلة بموضوع الاختبار، وتم الحصول عليه في نفس الفترة الزمنية تقريبًا، ويستخدم معامل الارتباط بين الدرجات الكلية للطلاب في الاختبار النهائي ودرجاتهم في أحد أدوات التقييم الأخرى في المقرر، مثل درجاتهم في أخر تكليف أو درجاتهم في مشروع للمقرر، ويراعى أن تكون هذه الأدوات ذات نتائج دقيقة وموثوقة، ويتحقق الصدق التلازمي للاختبار عندما يكون معامل الارتباط المشار إليه قوياً موجباً.

❖ خطوات حساب الصدق التلازمي

١- تحديد المحك التلازمي المناسب لقياس صدق الاختبار.

٢- إعداد قائمة بدرجات الطلاب على المحك المحدد.

٣- رصد درجات الطلاب في الاختبار التحصيلي التحريري في نفس القائمة السابقة.

٤- حساب معامل الارتباط بين درجات الطلاب في المحك ودرجاتهم في الاختبار التحريري.

وذلك كما هو موضح في الجدول التالي:

حساب معامل الارتباط

م	اسم الطالب	الدرجة في الاختبار التحريري (٤٠)	الدرجة في لمحك التلازمي (٢٠)
١		٣٥	١٨
٢		٣٢	١٧
٣		٣٢	١٥
٤		٣٦	١٩
٥		٢٨	١١
.	.....	....	....

❖ حساب الصدق التنبؤي للاختبار

يُعد الصدق التنبؤي دليلاً على أن درجة الاختبار ترتبط بمتغير سوف يتم تقييمه في مرحلة لاحقة لإجراء الاختبار. على سبيل المثال، حساب الارتباط بين درجات الطلاب في الاختبار ودرجاتهم في اختبار أخر لمقرر مشابه في مستوى دراسي أعلى، أو العلاقة بين درجات طلاب كلية التربية في مقرر استراتيجيات التدريس، وتقييم مدراء المدارس لهم مستقبلاً. ومن الجدير بالذكر أن معامل الارتباط في الصدق التنبؤي يكون أقل منه في الصدق التلازمي نتيجة وجود فاصل زمني أطول بين القياسين في الصدق التنبؤي، وهو ما يؤدي إلى تأثير عامل النمو والعديد من المتغيرات خلال هذه الفترة الزمنية على أداء العينة في القياس المحكي.

❖ خطوات حساب الصدق التنبؤي للاختبار

١. تحديد المحك التنبؤي المناسب لقياس صدق الاختبار.
٢. إعداد قائمة بدرجات الطلاب على المحك المحدد.
٣. رصد درجات الطلاب في الاختبار التحصيلي التحريري في نفس القائمة السابقة.
٤. حساب معامل الارتباط بين درجات الطلاب في المحك ودرجاتهم في الاختبار التحريري.

وذلك كما هو موضح في الجدول التالي

حساب معامل الارتباط

م	اسم الطالب	الدرجة في الاختبار التحريري (٤٠)	الدرجة في المحك التنبؤي (١٠٠)
١		٣٥	٩٠
٢		٣٢	٨٥
٣		٣٢	٨٣
٤		٣٦	٩٠
٥		٢٨	٧٠
...	.....	....	....

❖ ثبات الاختبار التحصيلي Reliability

يُقصد به اتساق واستقرار نتائج الاختبار بحيث يعزو التباين والفروق بين الطلاب في الدرجات إلى فروق حقيقية في التحصيل في المقرر وليس إلى عوامل أخرى، مما يؤدي إلى احتفاظ الطلاب بالموقع النسبي عبر القياسات المتكررة باستخدام نفس الاختبار أو صورة مكافئة له، بشرط تكرار نفس ظروف التطبيق.

مثال: يوضح الجدول التالي درجات أربعة طلاب على اختبار تم تطبيقه مرتين بفاصل زمني شهر

الطالب	درجات التطبيق الأول	المركز النسبي (الترتيب)	درجات التطبيق الثاني	المركز النسبي (الترتيب)
١	٩٣	١	٩٥	١
٢	٨٣	٢	٨٩	٢
٣	٧٨	٣	٨٠	٣
٤	٥٩	٤	٥٧	٤

تبين درجات إعادة الاختبار زيادة درجات الطلاب، فيما عدا الطالب رقم (٤)، إلا أن الاختلاف في الدرجات كان طفيفًا ولم يؤد إلى تغير ترتيب الطلاب، مما يشير إلى ثبات الاختبار.

وحيث أن درجة الفرد في الاختبار تتأثر بأخطاء القياس المختلفة، فإن **معامل ثبات الاختبار يُعبر عن مدى اقتراب هذه الدرجة من درجته الحقيقية**، ومن أمثلة مصادر أخطاء القياس: الغش، والتخمين، وقلق الاختبار، والعوامل البيئية غير المناسبة أثناء تطبيق الاختبار، وعدم تجانس محتوى الاختبار، ويُعبر عن العلاقة بين الدرجة الملاحظة للطالب ودرجته الحقيقية من خلال المعادلة التالية:

درجة الفرد الملاحظة (الدرجة الفعلية) = الدرجة الحقيقية ± درجة الخطأ في القياس.

طرق حساب ثبات الاختبار

❖ حساب معامل الثبات بطريقة إعادة التطبيق Test-Retest Reliability

يعبر هذا المعامل عن مدى اتساق أداء الطلاب في تطبيقين لنفس الاختبار بينهما فاصل زمني مناسب لا يسمح بالتذكر أو بتأثير عوامل النمو على أداء المختبرين، حيث يتم تطبيق نفس الاختبار على نفس الطلاب، وفيما يلي خطوات حساب معامل ثبات إعادة التطبيق:

١. تحديد الفاصل الزمني المناسب بين التطبيق، ومن ثم تحديد موعدا التطبيق.
٢. تطبيق الاختبار للمرة الأولى، وتصحيحه، ورصد درجته لكل طالب في قائمة معدة لذلك.
٣. تطبيق الاختبار للمرة الثانية على نفس الطلاب، واتباع نفس إجراءات التطبيق الأول.
٤. تصحيح الاختبار ورصد درجته لكل طالب في نفس القائمة السابقة.
٥. حساب معامل الارتباط بين درجات الطلاب في كلا التطبيقين، كما هو موضح في الجدول التالي:

حساب معامل الارتباط

م	اسم الطالب	الدرجة في التطبيق الأول (٤٠)	الدرجة في التطبيق الأول (٤٠)
١		٣٥	٣٤
٢		٣٢	٣٢
٣		٣٢	٣٠
٤		٣٦	٣٧
٥		٢٨	٢٩
...	....	....	....

❖ حساب معامل ثبات الصور المتكافئة Equivalent- Forms Reliability

تُعد الاختبارات المتكافئة ذات أهمية عملية في سياق التعليم الجامعي، حيث تُستخدم الصور المتكافئة في اختبارات الإعادة للطلاب المقبولة أعذارهم، أو عند إعادة الاختبار للطلاب لأي ظرف آخر، كما تستخدم الصور المتكافئة للحد من الغش عند زيادة كثافة الممتحنين في قاعات الاختبارات؛ ولذا يجب التحقق من تكافؤ الصور المختلفة للاختبار، وذلك

من خلال حساب معامل ثبات الصور المتكافئة، والذي يُعبر عن اتساق درجات الطلاب في نسختين متكافئين من الاختبار، ويتم حساب هذا المعامل من خلال الخطوات التالية (شكل٣):

١. إعداد صورتين متكافئتين في المحتوى من الاختبار، وذلك بناءً على جدول المواصفات، بحيث تتحقق شروط التكافؤ التي سبق ذكرها (يرجع إلى دليل بنوك الأسئلة والصورة المتكافئة الصادر عن مركز جودة التقويم والاختبارات).

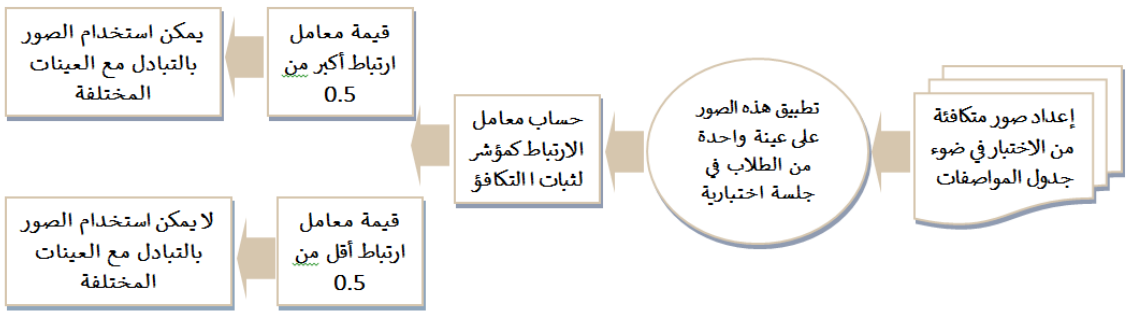
٢. تطبيق الصورتين في جلسة اختبارية واحدة بشكل متتابع على الطلاب مع مراعاة مناسبة طول الاختبارات لزمن التطبيق، وضمان عدم اختلاط الطلاب في الفترة بين الاختبارين لضبط عوامل الغش، بما لا يؤثر في درجات الطلاب عبر الاختبارين.

٣. تصحيح الصورتين ورصد درجات الطلاب.

٤. حساب معامل الارتباط بين درجات الطلاب على الصورتين، كما هو موضح في الجدول التالي:

حساب معامل الارتباط

م	اسم الطالب	الدرجة في الصورة الأولى (٤٠)	الدرجة في الصورة الثانية (٤٠)
١		٣٥	٣٤
٢		٣٢	٣٢
٣		٣٢	٣٠
٤		٣٦	٣٧
٥		٢٨	٢٩
...	.....	....	....



شكل٣. خطوات التحقق من تكافؤ الصور المختلفة للاختبار ثم استخدامها.

❖ معامل ثبات الاتساق الداخلي Internal Consistency Reliability

في بعض الأحيان يكون من الصعب بالنسبة لأستاذ المقرر إعداد صورتين متكافئتين من الاختبار، أو قد يصعب عليه تطبيق الاختبار التحصيلي نفسه مرتين متتاليتين لقياس ما تعلمه الطلاب، وفي هذه الحالات يُفضل تقدير ثبات درجات الاختبار باستخدام معامل الاتساق الداخليInternal Consistency، والذي يعتمد على استخدام اختبار واحد وتطبيقه مرة واحدة، ومن أمثلة معاملات الاتساق الداخلي التجزئة النصفية وألفا كرونباخ، وفيما يلي شرح كل منهما:

❖ معامل ثبات التجزئة النصفية Split-Half Reliability

تماثل طريقة حساب معامل ثبات التجزئة النصفية طريقة الصور المتكافئة، إلا أن في طريقة حساب معامل ثبات التجزئة النصفية يتم استخدام اختبار واحد وتجزئة الاختبار إلى نصفين متكافئين، من خلال الخطوات التالية:

١. إعداد صورتين متكافئتين للاختبار.
٢. دمج الصورتين المتكافئتين في ورقة اختبارية واحدة؛ أي اختبار واحد، بحيث تمثل المفردات الفردية الصورة الأولى، وتمثل المفردات الزوجية الصورة الثانية.
٣. تطبيق الاختبار على عينة واحدة من الطلاب.
٤. حساب درجة كل طالب في الأسئلة الفردية (درجته في نصف الاختبار)، والتي تمثل درجته في الصورة الأولى للاختبار.
٥. حساب درجة كل طالب في الأسئلة الزوجية (درجته في النصف الثاني للاختبار)، والتي تمثل درجته في الصورة الثانية للاختبار.
٦. حساب معامل ارتباط بيرسون بين درجات الطلاب على هذين النصفين. والذي يمثل في هذه الحالة معامل ارتباط نصف الاختبار.
٧. تصحيح معامل الارتباط الذي تم الحصول عليه في الخطوة السابقة للحصول على معامل ثبات درجات الاختبار ككل، وذلك باستخدام معادلة سبيرمان وبراون والموضحة أدناه:

معامل ثبات درجات الاختبار ككل

$$= \frac{2 \times \text{قيمة معامل ثبات درجات نصف الاختبار}}{1 + \text{قيمة معامل ثبات درجات نصف الاختبار}}$$

ولا تناسب طريقة التجزئة النصفية اختبارات السرعة، والتي تتطلب أن يجيب الفرد عن عدد كبير من الأسئلة السهلة نسبياً في زمن محدد. كما أنها تتطلب الدقة في تجزئة الاختبار بحيث يكون النصفين متكافئين.

❖ معامل ألفا كرونباخ لثبات الاتساق الداخلي Cronbach's Alpha

يُمثل معامل ألفا كرونباخ أحد الأدلة على تجانس محتوى المفردات الاختبارية واتساق أداء الطلاب عليها، كما يُعطي الحد الأدنى للقيمة التقديرية لمعامل ثبات درجات الاختبار، أي أن قيمة معامل ثبات درجات الاختبار عامة لا تقل عن قيمة معامل ألفا كرونباخ، ويُعد معامل ألفا كرونباخ أحد مؤشرات ثبات درجات الاختبار الأكثر شيوعاً واستخداماً، وذلك للأسباب التالية:

- ✓ سهولة الحصول على هذا المعامل من خلال تقارير التصحيح الآلي، أو باستخدام العديد من البرامج الإحصائية.
- ✓ الحصول عليه من خلال تطبيق الاختبار لمرة واحدة على عينة كافية من الممتحنين، ولا يتطلب حسابه إعداد صور متكافئة أو إعادة تطبيق الاختبار.

- ✓ يمثل معامل ألفا كرونباخ قيمة تقديرية لمتوسط قيم معامل الارتباط بين كلا من نصفي الاختبار دون إجراء هذه التجزئة فعلياً.
- ✓ يتراوح مدى معامل الثبات من صفر إلى 1.
- ✓ تتفاوت القيمة المقبولة لمعامل الثبات وفقًا للغرض من القرارات المستندة إلى نتائج الاختبار (يرجع إلى دليل تفسير واستثمار نتائج التحليل الإحصائي للاختبارات التحريرية الصادر عن مركز جودة التقويم والاختبارات).

العلاقة بين الثبات والصدق

إن الاختبار الذي يفتقر إلى الصدق يفتقر أيضًا إلى الثبات، والعكس ليس بالضروري أن يكون صحيحًا، فعلى سبيل المثال: الاختبار عالي الثبات قد لا يكون صادقا، إذ قد يعطي نفس النتائج عند إعادة التطبيق رغم أنه يقيس قدرة تختلف عن تلك التي ينبغي وأن يقيسها، كأن يتشبع اختبار تحصيلي في الإحصاء بالقدرات اللغوية.

الموضوعية Objectivity

يصبح الاختبار موضوعيًا إذا كانت النتائج مستقلة عن الأحكام الذاتية للمصحح، وتُعني موضوعية التصحيح عدم التأثر بذاتية المصحح، أي عدم اختلاف درجات الاختبار باختلاف المصححين، كما تعني أيضاً أن يكون الجواب محددًا سلفًا بحيث لا يختلف عليه اثنان، كما هو الحال في الأسئلة الموضوعية.

ويمكن حساب الموضوعية بحساب ثبات تقديرات المصححين، من خلال وضع مصفوفة للتصحيح rubric، ويتم تدريب المصححين عليها، ثم يتم حساب معامل الارتباط بين درجات المقدرين interrater- reliability لعينة من الأوراق الاختبارية لا تقل عن ١٠٪ من إجمالي الأوراق الاختبارية.

وتؤثر الموضوعية في ثبات نتائج الاختبار ويمكن تحسينها من خلال:

- ✓ مراعاة معايير جودة صياغة الأسئلة.
- ✓ وجود نموذج إجابة لتصحيح الأسئلة الموضوعية، ومصفوفة تقدير لتصحيح الأسئلة المقالية.
- ✓ مراجعة الاختبارات المصححة إلكترونيًا.
- ✓ قراءة عينة من الإجابات قبل التصحيح.
- ✓ عدم تأثر التصحيح بعوامل خارجية كخط الطالب، وأسلوبه، وتنسيق كتابته، وجودة إملاءه.
- ✓ تخلص المصححين من تأثير الهالة، وإصدار الأحكام المسبقة (هذا الطالب عالي المستوى، أو هذا الطالب منخفض المستوى)، وذلك قبل القيام بعملية التصحيح.

التوازن Balance

ويقصد بتوازن الاختبار أن تكون نسبة الأسئلة أو المفردات المتعلقة بمخرجات التعلم، وموضوعات المقرر متفقة مع النسب المحددة لها في جدول المواصفات، وبما يحقق شموليته، ويضمن صدق الاختبار، وثباته.

العدالة Fairness

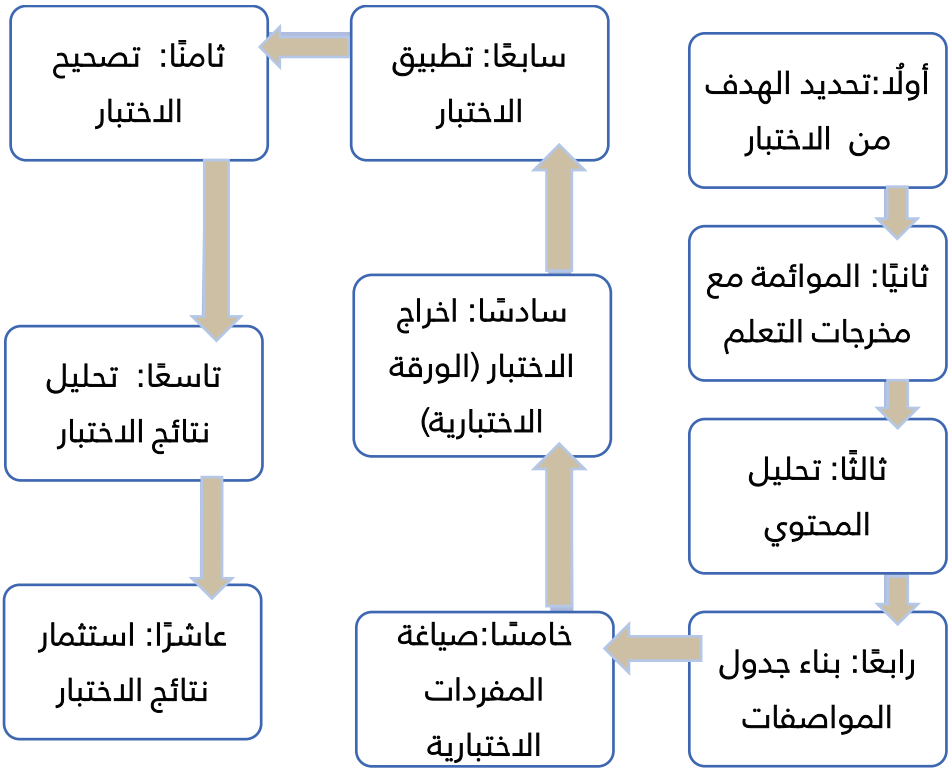
ويُقصد بالعدالة تمكين كل الطلاب العاديين وذوي الإعاقة والحالات الخاصة من إظهار ما أنجزوه من مخرجات التعلم، ويتم تحقيقها في جميع مرحل إعداد الاختبار بداية من صياغة مفرداته في ضوء جدول المواصفات، ثم الحيادية في بناء الأسئلة، وتجنب استخدام أمثلة متحيزة لفئة معينة، كأن تكون مستقاة من بيئة صناعية أو ريفية غير متاحة لجميع الطلاب، أو استخدام أمثلة متحيزة للنوع مثل المشكلات القائمة على ممارسة رياضة معينة يمارسها الذكور دون الإناث، ومرورًا بطريق تطبيق الاختبار في ظروف اختبارية مناسبة للجميع، وانتهاء بإتاحة الفرصة للطالب لأن يتظلم على درجاته إذا ما اعتقد أن درجته لا تعكس الأداء الحقيقي لقدراته، بالإضافة إلى تقديم الترتيبات اللازمة لمواءمة تقييم الطلاب ذوي الإعاقة والحالات الخاصة(يرجع إلى دليل مواءمة تقييم الطلاب ذوي الإعاقة والحالات الخاصة الصادر عن مركز جودة التقويم والاختبارات)

صلاحية الاستخدام Usability

ويقصد بها مدى صلاحية الاختبار للاستخدام الفعلي وسهولة ذلك بحيث يكون عمليًا واقتصاديًا في الإعداد، والتطبيق، والإجابة، والتصحيح، وتوفير الجهد والوقت، كما يجب أن يحتوي الاختبار على أدلة كافية تساعد في الحصول على أكبر قدر من المعلومات عن نسب إنجاز الطلاب لمخرجات التعلم.

خطوات إعداد الاختبار التحصيلي الجيد

تستهدف البرامج الأكاديمية التعرف على نسب إنجاز مخرجات التعلم من قبل الطلاب؛ مما يتطلب بناء اختبارات تحصيلية تستند إلى مخرجات التعلم، وتستهدف قياسها، وذلك من خلال اتباع الخطوات العلمية لإعداد الاختبار التحصيلي الجيد، والتي يمكن تلخيصها في الشكل التالي.



شكل٤.خطوات إعداد الاختبار التحصيلي الجيد

وفيما يلي شرح لهذه الخطوات بالتفصيل:

أولًا: تحديد الهدف من الاختبار

يجب تحديد الغرض من الاختبار قبل البدء في بنائه - كما سبق توضيح أغراض الاختبارات- وهل يهدف إلى متابعة مدى تقدم المتعلمين في المقرر (تقويم تكويني)، أم أنه اختبار تشخيصي لتحديد الصعوبات أو المفاهيم الخطأ لدى الطلاب، أم أنه بهدف التعرف على ما أنجزه الطلاب من مخرجات التعلم (نهائي).

ثانيًا: الموائمة مع مخرجات التعلم المراد قياسها

إن موائمة مفردات الاختبار مع مخرجات التعلم المستهدفة هو المفتاح إلى تصميم اختبار تحصيلي جيد، حيث تبدأ عملية التدريس بتحديد مخرجات التعلم والأهداف التعليمية، بمعنى تحديد ما يجب أن يحققه الطالب بنهاية عملية التعليم والتعلم، ووفقًا لذلك يتم عمل ملاءمة المحتوى التعليمي لمخرجات التعلم وآليات التنفيذ (الأساليب والأنشطة التعليمية، والإجراءات التعليمية الملائمة والميسرة لتحقيق الأهداف، ثم الإجراءات التي من شأنها الوقوف على فاعلية عملية التدريس وطرق التقويم وأدواته)، وذلك طبقاً لما أدرجه المركز الوطني للتقويم والاعتماد الأكاديمي (مركز اعتماد) التابع لهيئة تقويم التعليم والتدريب ضمن المعيار الثالث (التعلم والتعليم). ويتم تحديد مخرجات التعلم من خلال عبارات سلوكية تدل على الأداء النهائي المتوقع من الطالب. **ولكن ماهي العبارة السلوكية وكيف يتم صياغتها؟**

❖ العبارة السلوكية

يشير مفهوم العبارة السلوكية إلى وصف المخرج التعليمي المرغوب تحقيقه / الأداء النهائي من قبل الطلاب أو التغير الحادث في سلوك الفرد نتيجة التعلم، وينبغي أن يكون هذا السلوك قابل للملاحظة، وتتضمن العبارة السلوكية مجموعة من المكونات **نوضحها فيما يأتي:**

❖ المخرج (السلوك) النهائي

ويعبر عنه باستخدام أفعال سلوكية action verbs يتم اختيارها وصياغتها بدقة، ويتم اشتقاقها من الأهداف الرئيسة للتعلم، ويجب أن يكون السلوك المصاغ قابل للقياس والملاحظة، كما ينبغي برهنته من قبل الطالب، وأن يقترن بالمحتوى التعليمي (أن يكون له مرجعية).

❖ المستوى المقبول

ويُعرف بالمعيار أو المحك أو الأداء المطلوب لتحقيق مخرج التعلم، وقد يكون المستوى المقبول هو زمن الأداء من قبل الطلاب، أو دقة الإجابة، أو الكفاية التعليمية المطلوبة، أو المستوى المطلوب، فعلى سبيل المثال: قد يكون المستوى هو أداء التجربة بإتقان وبدون أخطاء، أو بحد أقصى للأخطاء والذي إذا تم تجاوزه لا يصنف المتعلم على أنه حقق مخرج التعلم، وقد يكون محكًا (٨٠٪) مثلاً؛ ويعنى وجود المستوى أن مخرج التعلم قابل للقياس والمقارنة وإظهار الفروق الفردية بين الطلاب.

❖ شروط مخرج التعلم

وتعني الشروط، أو الظروف، أو المعطيات أو المصادر التي يُسمح للطالب باستخدامها عند القيام بالأداء، والتي يجب أن يتم تحقيق الهدف تحت تأثيرها أو في حالة توفرها، وعلى سبيل المثال: حل المعادلات مستعينًا بالآلة الحاسبة، أو ترجمة فقرة محددة باستخدام القاموس، أو الوصول لاستنتاج محدد في المعمل، أو القيام بالسلوك في الظروف التقليدية دون الحاجة إلى أي مصدر أخر .... إلخ.

وعليه يكون مكونات صياغة العبارة السلوكية:

- ✓ أن المصدرية: أن.....
- ✓ الفعل الدال على السلوك/ المخرج: يتعرف، يكتب، يناقش، يقارن، يعلل.....
- ✓ القائم بالسلوك وهو الطالب/ المتعلم
- ✓ معيار الأداء: المستوى المقبول للأداء في ضوء محكات.
- ✓ الشروط أو محددات الأداء: السياق الذي يظهر خلاله الأداء المطلوب.

أمثلة:

- أن يقرأ المتعلم الدرس قراءة جهرية بدون أخطاء نحوية في مجموعته.
- أن يميز المتعلم بين مفهومي الصدق والثبات للاختبار بدقة من خلال الرسوم التوضيحية.
- أن يكتب الطالب مقالًا عن النزاهة باللغة الإنجليزية الصحيحة أثناء المحاضرة وبالتعاون مع مجموعته.
- أن يحدد الطالب على خريطة صماء الفتوحات الإسلامية في عصر الخلفاء الراشدين بدقة لا تقل عن ٩٠٪ من خلال الواجب المنزلي.

وقد لا تتطلب صياغة مخرجات التعلم وفق NCAAA التوضيح التفصيلي للعبارة السلوكية بكل أركانها نظرًا لأنها تُصاغ على مستوى المقرر، وفي حال قيام أستاذ المقرر ببناء مخرجات للتعلم على مستوى المحاضرة الواحدة؛ فيجب اكتمال كافة أركان العبارة السلوكية.

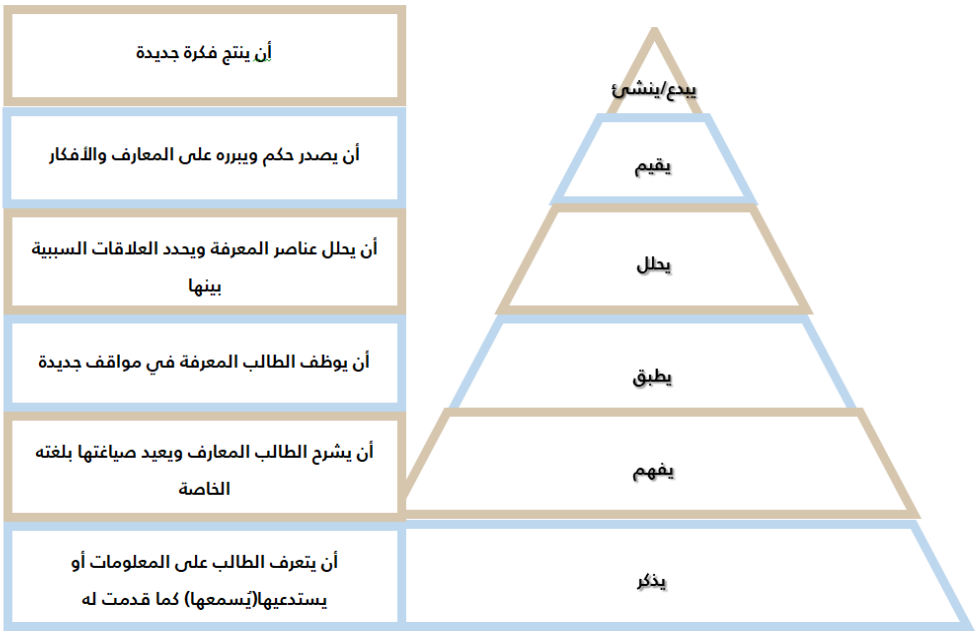
❖ تصنيف مخرجات التعلم إلى مجالات ومستويات

تم تصنيف مخرجات التعلم في ثلاث مجالات طبقاً للإطار الوطني للمؤهلات (٢٠٢٠، ص١٣) هي المعرفة والفهم، والمهارات، والقيم، كالتالي:

المعرفة والفهم	المهارات	القيم
وتشمل ما يعرفه الخريج ويفهمه في مجال التعلم والعمل أو المهنة، من: • المعرفة العميقة الواسعة، وفهم الحقائق والمفاهيم والمبادئ والنظريات والعمليات والإجراءات المتضمنة في مجال التعلم أو العمل أو المهنة . • عمق المعرفة، حيث يمكن أن يكون عاما ً أو متخصصاً. • اتساع مجال المعرفة، حيث يتراوح بين موضوع واحد ومتعدد التخصصات .	وتشمل ما يستطيع الخريج القيام به للتعلم المستمر والعمل أو المهنة، من: • المهارات الإدراكية: وتتضمن تطبيق المعرفة، ومهارات التفكير الناقد وحل المشكلات، والاستقصاء، والإبداع . • المهارات العملية والبدنية: وتتضمن تطبيق المعرفة، واستخدام المواد والأجهزة والأدوات المناسبة، وتطبيق المهارات الحركية واليدوية ببراعة	وتشمل ما يتمثل له الخريج من مبادئ ومعايير توجهه للنجاح بمجالات الحياة والعمل أو المهنة، وتشمل • القيم والأخلاق الأكاديمية والمهنية . • التعلم الذاتي المستمر والاستقلالية . • العمل مع فريق وتحمل المسؤولية.

المعرفة والفهم	المهارات	القيم
• أنواع المعرفة، وتتدرج من معرفة حسية، إلى معرفة مجردة، ومن معرفة بسيطة إلى معرفة تراكمية . • تعقيد المعرفة، حيث يجمع بين نوع المعرفة، وعمقها، واتساعها.	• .مهارات التواصل وتقنية المعلومات :وتتضمن التواصل الكتابي واللفظي وغير اللفظي، والمهارات العددية، واستخدام تقنية المعلومات والاتصال وإنتاجها.	

سوف نقتصر في هذا الدليل على المجال المعرفي والمهارات الإدراكية فقط، وذلك لما يمثله من مخرجات تعليمية يمكن قياسها باستخدام الاختبارات التحصيلية موضوع الدليل.، يصنف بلوم (النموذج المعدل، ا . . ٢) المجال المعرفي إلي ست مستويات تترتب ترتيبًا هرميًا يبدأ بالمستوى الأبسط تفكيرًا وهو التذكر وينتهي بالمستوى الأكثر تعقيدًا وهو التركيب/ الإبداع، كما هو موضح بالشكل **أدناه:**



شكل5. المستويات المعرفية وفق نموذج بلوم المعدل2001

وفيما يلي نوضح أهم الأفعال السلوكية المستخدمة في كل مستوى على حدة:

مجال التقييم	المستوى المعرفي		أمثلة الأفعال السلوكية
1-المعرفة والفهم	التذكر	مهارات التفكير الدنيا وتتضمن معرفة الحقائق والقوانين والمصطلحات والتواريخ.....إلخ	يذكر، يسمي، يتلو، يحدد، يتعرف على، يصف، يعدد، يتعرف، يكتب قائمة، يصف، يسمي، يرتب، ينسخ، يربط، يضع في قائمة، يكرر.
	الفهم	معرفة" ماذا؟"	يوضّح، يفسّر، يناقش، يصيغ بأسلوبه، يعيد ترتيب، يستنبط، يستنتج، يلخّص، يترجم، يقارن، يصف، يشرح، يصف، يصنف، يحول، يقدر، يستدل على، يدلل، يعبر عن، يتنبأ، يعيد ترتيب، يراجع، يعيد كتابة، يكتب تقرير، يعدل، يلخص، يعيد الصياغة، يعلل.
٢-المهارات الإدراكية (مهارات التفكير العليا)	تطبيق	معرفة "كيف؟"	يطبق، يستخدم، يعلل، يحل (مسألة، أو مشكلة)، يحسب، يوضح بأمثلة، ينفذ، يختار، يكتشف، يوظف، يرتب، يكتب تقرير، يراجع، يظهر، يخطط، يحل مشكلة، يترجم، يجدول.
	تحليل		يحلل، يميز، يوازن، يقسم الموضوع، يقارن، يصنف، يفحص، يرسم بيانيّا، يختبر، يبرهن على صحة، يحلل موضوعاً إلى عناصره، يستنتج.
	تقييم		يصدر حكم –يقيم- يحقق- ينقد- يقوم، يقدر، يحكم، يفرض، يستنتج، يقارن، يفند، يحدد الصلة.
	إبداع		يقترح، يصمم، يبتكر، يؤلف، يعد، يربط، يخطط، يشكل، يكون، يعيد ترتيب، يعيد بناء، يركب، يخترع، ينشئ، يعيد تنظيم، يؤلف.

جدير بالذكر أن الفعل السلوكي، مرن ويمكن أن يُستخدم في أكثر من مستوى وفق السياق الذي يُستخدم فيه ونوع المعرفة التي يتم التعامل معها.

ثالثًا: تحليل المحتوى

ويُقصد به تحليل ما تم تدريسه في الوحدة من مفاهيم، ومصطلحات، وتعميمات، وحقائق، وعلاقات، ومهارات معرفية، وتحديد المستويات المعرفية المتضمنة في مخرجات التعلم.

رابعًا: بناء جدول المواصفات

❖ تعريف جدول المواصفات

هو مخطط لبناء الاختبار يُستخدم بهدف مواءمة الاختبار مع مخرجات التعلم من خلال انتقاء عينة ممثلة من جميع الأسئلة الممكنة في مقرر معين، ويتمثل في جدول يوضح العلاقة بين مخرجات التعلم المستهدفة وموضوعات المقرر، ويتم تحديد الأوزان النسبية لمخرجات التعلم المقاسة، وعليه يتم تحديد عدد الأسئلة التي ستقيس كل مخرج تعلم من خلال المحتوى الذي تم تدريسه، وذلك في مجالات تعلم المعرفة والفهم والمهارات الإدراكية وما تمثله من المستويات المعرفية للأهداف في تصنيف بلوم.

❖ أهمية جدول المواصفات

- ✓ مواءمة أسئلة الاختبار مع مخرجات التعلم المستهدفة من مجالات الفهم والمعرفة والمهارات الإدراكية.
- ✓ تحقيق الصدق الظاهري، وصدق المحتوى للاختبار من خلال انتقاء عينة ممثلة من الأسئلة التي تقيس مخرجات التعلم والموضوعات التي تم تدريسها.
- ✓ يساعد في بناء صور متكافئة للاختبار الواحد وبناء الاختبار البديل.
- ✓ التحقق من ملاءمة الاختبار لمخرجات التعلم وموضوعات المقرر.
- ✓ إكساب الطالب ثقة في عدالة وشفافية الاختبار، مما يساعد في تنظيم وقته أثناء الاستذكار.
- ✓ أحد أدوات ضمان جودة الاختبارات.

❖ نماذج جدول المواصفات ومكوناتها

تتعدد نماذج جدول المواصفات، وأكثرها شيوعًا هو النموذج ثنائي البعد (الموضوعات ومخرجات التعلم). وبناء على ذلك قد صمم مركز جودة التقويم والاختبارات بجامعة الإمام عبد الرحمن بن فيصل، نموذجين لجدول المواصفات هما نموذج (٢ - أ) والموسوم بمخطط الاختبار التحريري، ونموذج (٢ - ب) والموسوم بنموذج جدول مواصفات الاختبار التحريري (دليل سياسات وإجراءات التقويم والاختبارات، ٢٠٢١ ، ص ٨٤-٨٥)، ويعتمد نموذج (٢-أ) على تحديد الأوزان النسبية لمخرجات التعلم ومواءمة الموضوعات والأسئلة وفقًا لذلك، وبالتالي تحديد نوع و عدد الأسئلة التي تقيس كل مخرج تعلم، مما يُمكن الأستاذ من تصنيف أسئلة الاختبار في مجموعات، كل مجموعة تقيس مخرج تعلم محدد، وهذا بدوره يساعد على قياس مستوى إنجاز كل مخرج تعلم، وإعداد تقرير المقرر، وعليه يوصي مركز جودة التقويم والاختبارات الكليات بتبني نموذج (٢-أ)، وفيما يلي شرح هذا النموذج بالتفصيل.

❖ مكونات النموذج

يشتمل جدول المواصفات في ضوء هذا النموذج على:

- ✓ البيانات الأساسية للمقرر (الكلية - المسار- اسم المقرر- رمزه).
- ✓ البيانات الأساسية للاختبار (نوع الاختبار- الدرجات المخصصة - التاريخ- قاعة الاختبار- شكل الاختبار).

- ✓ مخرجات التعلم كما تم تحديدها في توصيف المقرر، وينبغي التركيز على أن مخرجات التعلم التي يقيسها الاختبار التحريري تقتصر على مجالات المعرفة والفهم والمهارات الإدراكية، ويتم كتابة عبارة مخرج التعلم مما يسهل مراجعة دقة المواءمة بين مخرجات التعلم والأسئلة التي تقيسها.
- ✓ مؤشرات الأهمية النسبية لمخرجات التعلم (مدة التدريس- حجم المحتوى- مدى التأثير – مدى التكرار- الاختبارات السابقة).
- ✓ الوزن النسبي لمخرج التعلم في ضوء أكثر ثلاث مؤشرات أهمية (يتم تثبيت نفس المؤشرات لجميع مخرجات التعلم بالمقرر).
- ✓ المجموع الكلي للأوزان النسبية لمخرجات التعلم.
- ✓ الدرجات المخصصة لكل مخرج تعلم.
- ✓ المجموع الكلي لدرجات الاختبار التحريري.
- ✓ المفردات التي تقيس كل مخرج تعلم.

❖ خطوات ملء النموذج

١. **الاستناد إلى توصيف المقرر عند إعداد جدول المواصفات باستخدام نموذج (٢ – أ) وذلك لاستكمال البيانات التالية:**

- ✓ بيانات المقرر، والتي تشمل اسم الكلية، واسم المقرر، ورمزه.
- ✓ بيانات الاختبار، وذلك طبقاً لخطة التقييم المدرجة بالتوصيف التي تُحدد درجة كل من الاختبار القصير، والاختبار الفصلي، والاختبار النهائي.
- ✓ عبارات مخرجات التعلم في مجالات المعرفة والفهم، والمهارات الإدراكية، ويتم إدراج كل منها في الصفوف المخصصة لها في النموذج.

٢. **موائمة موضوعات المقرر مع مخرجات التعلم،** حيث يتم الرجوع لموضوعات المقرر المحددة بتوصيفه، وكتابة عنوان الموضوع أو الموضوعات التي يتحقق من خلالها مخرج التعلم وذلك في العمود الموسوم بـ "الموضوعات" في النموذج، وتُعد هذه الخطوة مرشداً للأستاذ عند صياغة المفردات الاختبارية، حيث يقوم بصياغة المفردة التي تقيس مخرج تعلم محدد باستخدام محتوى علمي محدد تم تدريسه للطلاب.

٣. **مناقشة مؤشرات أهمية مخرجات التعلم من قبل فريق التدريس بالمقرر وتبني ثلاثة فقط منها، بما يتفق مع طبيعة المقرر،** مع مراعاة أن تظل هذه المؤشرات ثابتة ومتسقة عبر كل مخرجات التعلم بجدول مواصفات المقرر الواحد، علمًا بأنه إذا كان هناك اختبار واحد فقط بالمقرر، فإن مؤشر الاختبارات السابقة لا ينبغي اختياره كمؤشر للأهمية. ويراعى أن جدول المواصفات يتم إعداده مرة واحدة للمقرر، ولا يتم تغييره إلا إذا تم تعديل التوصيف وتغيير مخرجات التعلم في التوصيف. ومع ذلك يمكن فقط تغيير الأسئلة بالعمود الأخير، وفيما يلي شرح المقصود بمؤشرات أهمية مخرجات التعلم، وطريقة تقدير أهميتها:

- ✓ **مؤشر مدة التدريس Duration** ويشير إلى طول المدة المستغرقة في تحقيق مخرج التعلم مقدرًا بالساعات التدريسية أو عدد المحاضرات، وعليه إذا كانت هذه المدة قصيرة تُقدر ب ١، وإذا كانت هذه المدة طويلة ب ٣، وما بينهما يُقدر ب ٢.

- ✓ **مؤشر حجم المحتوى Content** ويُشير إلى حجم المادة العلمية وتعقد المفاهيم التي يتحقق من خلالها مخرج التعلم، ويُستخدم التقدير السابق، حيث يّعبر ١ عن مخرج تعلم ذو محتوى ضعيف ويعبر ٣ عن المحتوى الكثيف، و٢ لما بينهما.
- ✓ **مؤشر مدى التأثير Impact** ويُشير إل مدى أهمية مخرج التعلم في مهنة الطالب في المستقبل، وفي أدائه المهني، وعلى سبيل المثال: مخرج التعلم الذي يتعلق بسلامة وأمان المريض يصنف على أنه بالغ التأثير في تخصصات المهن الصحية، ويستخدم التقدير السابق من الأقل تأثيرًا (١) إلى (٣) للأكثر أهمية وتأثيراً.
- ✓ **مؤشر مدى التكرار/ الشيوع Frequency**، ويشير إلى مدى تكرار استخدام الطالب وتوظيفه لمخرج التعلم في أدائه المهني، ويستخدم التقدير السابق من (١) إلى (٣) ليدل (١) على المخرج نادر الاستخدام في المهنة لأداء المهني، بينما (٣) يعبر عن المخرج الأكثر شيوعًا واستخدامًا في المهنة المستقبلية و (٢) لما بينهما.
- ✓ **مؤشر الاختبارات السابقة Past Exams** ويقدر فيه مخرج التعلم بـ (١) إذا سبق تضمينه في أي اختبار تحريري سابق، بينما يقدر ب (٢) إذا لم يسبق تضمينه في أي اختبار تحريري، ويجب استبعاد هذا المؤشر عندما تقتصر الاختبارات في المقرر على اختبار تحريري واحد فقط (نهائي في معظم الحالات).

٤. **حساب الوزن النسبي لكل مخرج تعلم لتحديد نسبة الدرجات المخصصة له من مجموع درجات الاختبار وذلك من خلال:**

أ - ضرب تقديرات مؤشرات الأهمية الثلاثة الأكثر ارتباطًا بمخرجات تعلم المقرر.

ب - جمع حواصل الضرب لكل مخرج للحصول على المجموع الكلي لمؤشرات مخرجات الأهمية.

ج - قسمة حاصل ضرب المؤشرات لكل مخرج على المجموع الكلي لمؤشرات مخرجات لأهمية.

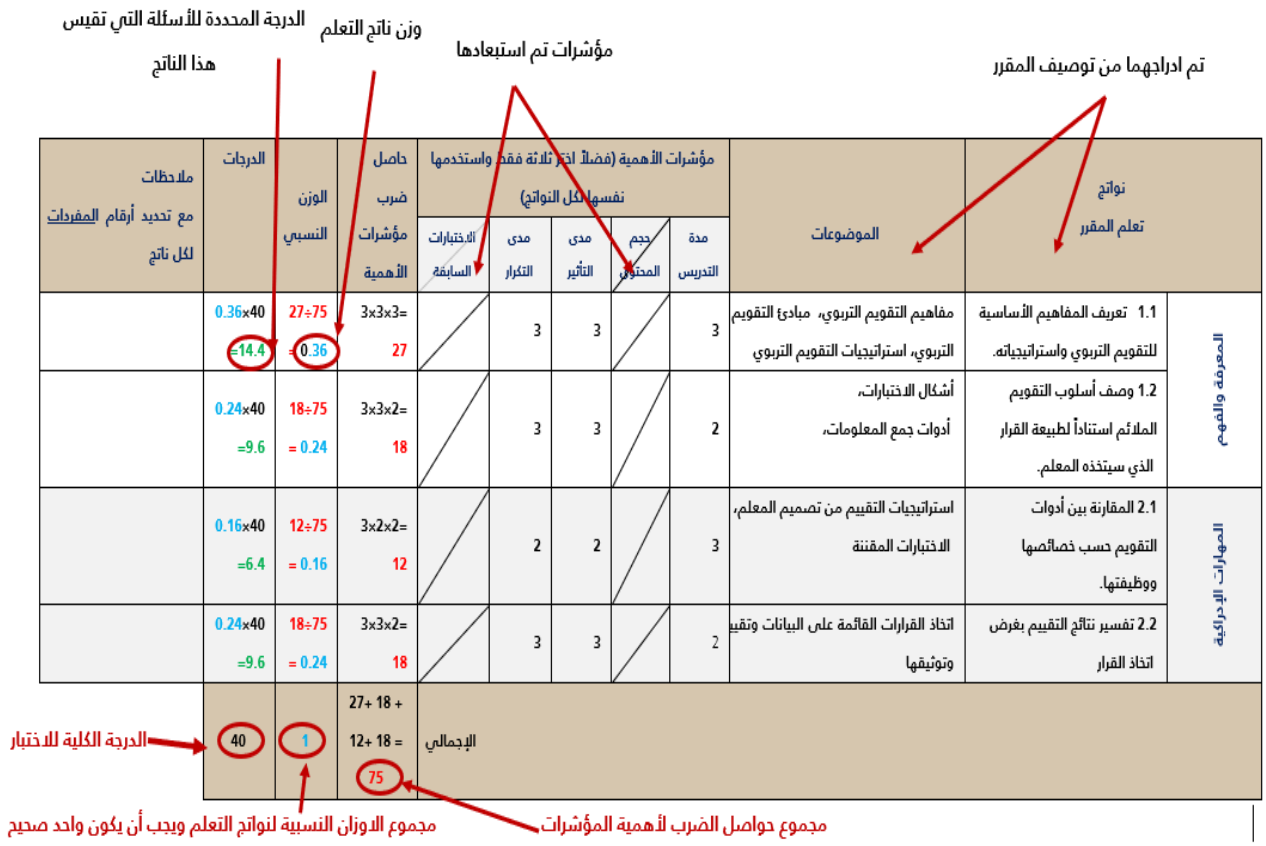
٥. **حساب عدد الدرجات المخصصة لكل مخرج تعلم من خلال ضرب الأهمية النسبية لكل مخرج في الدرجة الكلية للاختبار التحريري.**

٦. **إدراج أرقام المفردات التي تتواءم مع كل مخرج تعلم ويماثل مجموع درجاتها الدرجات المخصصة لكل مخرج تعلم في العمود الأخير من النموذج، مع تحديد نمط كل سؤال كما هو محدد بالاختبار.**

والخطوة الأخيرة (رقم ٦) لا تتم إلا بعد صياغة المفردات الاختبارية، وإخراج الاختبار ومراجعته، حيث يستخدم العمود الخاص بإدراج أرقام المفردات في المراجعة والتحقق من مواءمة أسئلة الاختبار مع مخرجات التعلم وفقًا لجدول مواصفات الاختبار، وعليه يتم التعديل عليه لكل اختبار على حدة. ويوضح النموذج التالي مثال لجدول مواصفات الاختبار النهائي في إحدى المقررات:

ويمكن الاستعانة بنموذج (٢ ـ أ) Excel الذي قام مركز جودة التقويم والاختبارات بتطويره على مستوى الاختبار التحريري الواحد، أو على مستوى جميع اختبارات المقرر، بحيث يتم إجراء حساب الأوزان النسبية والدرجات بشكل تلقائي عندما يقوم الأستاذ بإدراج الدرجة النهائية للاختبار، وتقدير مؤشرات الأهمية لكل مخرج تعلم (يمكن الحصول على النموذج من موقع الجامعة/ الخدمات الإلكترونية/ نماذج/ مركز جودة والاختبارات).

مثال: جدول مواصفات معبأ لأحد المقررات باستخدام نموذج ٢،أ



خامسًا: صياغة المفردات الاختبارية

- صياغة عدد من الأسئلة، أو ما تعرف بالمفردات الاختبارية بحيث يواءم نمط السؤال، والدرجة المخصصة له مع كل مخرج تعلم يقيسه، والوزن النسبي المقدر له (يرجع إلى دليل صياغة أسئلة الاختبارات التحريرية الصادر عن المركز).
- مناسبة الزمن المخصص للإجابة على كل نمط من أنماط الأسئلة مع الزمن الكلي للاختبار، **ويمكن الاستعانة بالجدول المدرج في ملف الإكسل الخاص بجدول المواصفات ٢-أ** لحساب الوقت المقدر لكل نمط من أنماط الأسئلة، وفيما يلي شرح هذا الجدول بالتفصيل:

١. إدراج عدد المفردات في كل نمط من أنماط الأسئلة التي تم تضمينها في الاختبار، ولا يشترط استخدام جميع أنماط المفردات المدرجة بالجدول.
٢. إدراج الزمن الذي تتطلبه المفردة في كل نمط من أنماط المفردات.
٣. حساب الزمن الكلي الذي يستغرقه كل نمط من أنماط الأسئلة، وذلك من خلال حاصل ضرب الزمن المطلوب للإجابة عن المفردة في عدد المفردات.
٤. إدراج الدرجة المخصصة لكل مفردة.

٥. حساب الدرجة المخصصة لكل نمط من أنماط الأسئلة من خلال ضرب درجة المفردة في عدد المفردات.

٦. حساب النسبة المئوية للدرجة المخصصة لكل نمط، وإدراجها في العمود الأخير من الجدول، وذلك للتحقق من عدم تجاوز درجات الأسئلة الموضوعية نسبة ٦٠٪ من إجمالي درجات الاختبار (يرجع للسياسات العامة لعملية جودة التقويم والاختبارات، ص٢٧).

مثال: جدول يوضح مراجعة دقة حساب الزمن الكلي للاختبار التحريري، كما هو مدرج في ملف الاكسل الخاص بنموذج أ،٢

جدول حساب النسبة المئوية والوقت المقدر لكل نوع من الأسئلة

	نمط السؤال	عدد المفردات	الوقت المقدر للمفردة	إجمالي الوقت المقدر لكل نوع	الدرجة المحددة للمفردة	إجمالي الدرجات المقدر لكل نوع	٪ الدرجة
أسئلة موضوعية	اختيار من متعدد						
	اختيار من متعدد لقياس مهارات التفكير العليا						
	صواب أو خطأ						
	صواب أو خطأ مركب						
	المزاوجة						
أسئلة مقالية	تكملة						
	إجابة قصيرة						
	مقالي						
	المجموع : الزمن المتوقع للاختبار				مجموع درجات الاختبار		١٠٠ %

تحقق من أن إجمالي الوقت المقدر لا يتجاوز الوقت المحدد للاختبار، مع الأخذ في الاعتبار الوقت الذي يستغرقه الطالب في استكمال بياناته ومراجعة الإجابة

تحقق من أن إجمالي الدرجات متطابق مع النهاية العظمى للاختبار

سادسًا: إخراج ورقة الاختبار

❖ تجميع الأسئلة وترتيبها

ينبغي ترتيب الأسئلة في الورقة الاختبارية في ضوء الإرشادات التالية:

- ✓ وضع مفردات الاختبار من نفس النمط، بحيث تُقدم تعليمات الإجابة الخاصة بكل نمط من أنماط الأسئلة مرة واحدة، مما يسهل عمليات التصحيح وتحليل النتائج، ويحافظ على التهيؤ العقلي للطالب.
- ✓ يفضل ترتيب المفردات وفقًا لدرجة السهولة؛ بحيث تظهر الأنماط السهلة في بداية الاختبار، مما يزيد من دافعية الطالب لأداء الاختبار، ويقلل من فرص الإحباط وتأثير قلق الاختبار الذي قد يظهر نتيجة البدء بمفردة صعبة.
- ✓ يفضل تجميع أنماط الأسئلة التي تقيس نفس مخرجات التعلم معا، وذلك لأغراض تدريسية تتعلق بتحديد مجالات التعلم التي يواجه الطلاب فيها صعوبات بغرض علاجها.

❖ تعليمات الاختبار

ينبغي صياغة تعليمات اختبار بسيطة ومختصرة وواضحة ومفهومة وتحتوي على المعلومات الآتية:

- ✓ غرض الاختبار.
- ✓ زمن الاختبار.
- ✓ كيفية تسجيل إجابات الاختبار في ورقة الإجابة.
- ✓ تعليمات للجوء إلى التخمين في حالة عدم معرفة الإجابة.

❖ مواصفات الورقة الاختبارية:

يتكون الاختبار التحصيلي من مجموعة من الأوراق التي يمكن تسميتها بالورقة الاختبارية أو كراسة الاختبار تشتمل على:

- ✓ ورقة الغلاف وتتضمن اسم الجامعة والكلية والقسم والبرنامج- نوع الاختبار- اسم المقرر ورمزه- تاريخ عقد الاختبار- المستوى والشعبة - زمن الاختبار- بيانات الطالب (اسمه ورقمه الجامعي) - أسماء أساتذة المقرر ومنسقه.
- ✓ تعليمات الاختبار.
- ✓ تنبيه الطلاب إلى قراءة التعليمات قبل الإجابة (إجابة السؤال- تدوين الإجابة في المكان المخصص)
- ✓ عدد أسئلة الاختبار.
- ✓ عدد صفحات الاختبار.

ويحتوي دليل سياسات التقويم وإجراءاته على نموذج قائمة تحقق للورقة الاختبارية الجيدة، ينبغي مراجعة الورقة الاختبارية في ضوئه، وتوقيع المراجع واعتماد رئيس القسم قبل طباعتها (نموذج ٤: التحقق من جودة الورقة الاختبارية بدليل سياسات وإجراءات التقويم والاختبارات).

سابعًا: تطبيق الاختبار

يتضمن تطبيق الاختبار توفير الظروف الاختبارية الملائمة لتأديته من قبل الطلاب، والتحكم في أي أحداث طارئة من شأنها أن تؤثر سلبًا في أدائهم، وتسهم في زيادة أخطاء القياس مما يؤثر في ثبات الاختبار، ولذلك يجب مراعاة ما يلي:

- ✓ خفض احتمالية أي حدوث ما يعطل الاختبار، وذلك باختيار غرفة مناسبة ومجهزة، مع توحيد ظروف تطبيق الاختبار في كل لجانه.
- ✓ وضع مسافات كافية بين الطلاب لتقليل محاولات الغش.
- ✓ عدم إشعار الطلاب بصعوبة الاختبار؛ تجنبًا لارتفاع مستوى القلق لديهم.
- ✓ وضوح التعليمات المكتوبة للاختبار، ويمكن توضيحها لفظيًا في حالات الضرورة، لكن ليس من قبل أستاذ المقرر.
- ✓ التحقق من مطابقة عدد الطلاب مع عدد أوراق الأسئلة والإجابة.
- ✓ رصد الغياب والحضور للمتقدمين للاختبار.
- ✓ التزام الطلاب بقواعد الانضباط داخل القاعة الاختبارية.
- ✓ عدم السماح بالغش أو التشجيع عليه.
- ✓ تنفيذ خطة طوارئ الاختبارات في حال إذا ما استدعى الأمر ذلك.

ثامنًا: تصحيح الاختبار

- ❖ يجب إعداد نموذج إجابة للاختبار يشمل مفرداته بنوعيها الموضوعية والمقالية، على أن يتضمن الإجابات الصحيحة والمقبولة، لتمثل محكًا لمدى وضوح أسئلة الاختبار ووجود إجابات محددة لها، كما يوضح نموذج الإجابة طرق أو خطوات حل الأسئلة وتوزيع الدرجات على خطوات الحل، كما يجب مراجعة نموذج الإجابة مصفوفة التقدير scoring rubric من قبل أستاذ في التخصص.

ويجب مراعاة ما يلي للوصول إلى أفضل مستوى من دقة **تصحيح الاختبار:**

- ✓ إعداد نموذج الإجابة عن السؤال عقب إعداده مباشرة، وتحديد درجة كل جزئية فيه.
- ✓ مطابقة نموذج الإجابة مع السؤال مطابقة تامة، حتى لا يخصص جزء من درجة السؤال ليشمل معلومة غير واردة بوضوح في نص السؤال، فالسؤال ينبغي أن يقيس تحصيل الطالب وليس ذكائه أو درجة انتباهه.

- ✓ قراءة عينة من الإجابات قبل التصحيح، بحيث يمكن في ضوئها تعديل توزيع درجة السؤال على جزئياته، أو تعديل نموذج الإجابة ليتضمن إجابات لم يتضمنها الأصل.
- ✓ التزام المصحح بالإجابة النموذجية للاختبارات الموضوعية، ومصفوفة التقدير للاختبارات المقالية.
- ✓ تعميم كل قرار أو إجراء يتم اتخاذه في حالة وجود أكثر من مصحح للاختبار، وخاصة عندما يتم تعديل على الإجابة النموذجية، أو وجود إجابة بديلة أخرى.

❖ **يراعى في الأسئلة المقالية التالي:**

- ✓ تصحيح سؤال واحد في جميع أوراق الإجابة حتى ينتهي تصحيح نفس السؤال تمامًا، ثم يتم الانتقال إلى تصحيح سؤال ثان بنفس الطريقة، مما يُقلل من أثر الهالة، أو تأثر تقدير الإجابة في السؤال التالي بدرجات الطالب في السؤال السابق له خاصة، وإن كانت مرتفعة، مما قد يؤدي إل خطأ في القياس، ومنح إجابة السؤال التالي درجات أكثر مما تستحق.
- ✓ تصحيح إجابات السؤال الواحد في جميع الأوراق جلسة واحدة لضمان اتساق التقدير.
- ✓ ترتيب الأوراق الاختبارية بطريقة عشوائية بعد تصحيح كل سؤال، لتجنب تأثير درجة الطالب بدرجة الطالب الذي يسبقه.
- ✓ يجب ألا يتأثر التصحيح بعوامل خارجية كخط الطالب، أو أسلوبه، أو ترتيبه، أو جودة إملاءه، أو علامات الترقيم، بالإضافة إلى أثر الهالة التي تحيط به، ويمكن تخصيص درجات مستقلة لبعض العوامل مثل سلامة اللغة مثلًا إذا كان ذلك أحد المتطلبات في مواصفات الخريج للمنافسة في سوق العمل.
- ✓ ينبغي عدم معرفة اسم الطالب الذي يقوم المصحح بتصحيح إجابته، لتجنب التحيز وضمان الموضوعية.
- ✓ التحقق من دقة التصحيح الإلكتروني وذلك من خلال مراجعة عينة عشوائية من أوراق الإجابة المصححة إلكترونيًا، وهو أمر مهم لضمان سلامة نتائج تحليل المفردات.

تاسعًا: تحليل نتائج الاختبار

تهدف هذه المرحلة إلى التحقق من جودة المفردات الاختبارية، صدق وثبات الاختبار، ويساعد التصحيح الآلي للاختبارات في إنجاز هذه المرحلة بسهولة، حيث يمكن للأساتذة استخراج مجموعة شاملة ومتنوعة من التقارير التحليلية عن أداء الممتحنين، الاختبار، المفردات، ومخرجات التعلم (يرجع إلى دليل تحليل واستثمار نتائج الاختبارات التحريرية الصادر عن مركز جودة التقويم والاختبارات).

عاشرًا: استثمار نتائج الاختبار

تهدف هذه المرحلة إلى الاستفادة من التقارير المستخرجة في المرحلة السابقة لضمان تعزيز جودة الاختبارات التحريرية، وإغلاق دائرة التقييم، ودعم عمليات التعليم والتعلم على مستوى البرنامج الأكاديمي (يرجع إلى دليل تحليل واستثمار نتائج الاختبارات التحريرية الصادر عن مركز جودة التقويم والاختبارات).

الاختبارات الشفوية

تعتمد الاختبارات التحريرية وبخاصة المقالية اعتمادًا أساسيًا على مهارات الطلاب القرائية والكتابية literacy skills، وعليه قد تتأثر نتيجة الطالب في الاختبار بقدراته التعبيرية، مما قد يؤدي إلى خطأ في القياس، أو أن يتأثر المصحح بأثر الهالة للطالب بمعنى أن يعطى الطالب المتميز في الكتابة انطباعًا جيدًا عن أدائه التحصيلي في أسئلة المقال، حتى وإن كان مفتقرًا إلى المعرفة التي تقيسها الأسئلة؛ وهو ما يؤدي إلى وجود اختلاف في تقييم الطالب الضعيف في تلك المهارات، أو الطالب ذوي الصعوبات في القراءة أو الهجاء، أو الطالب الضعيف في اللغة الإنجليزية في مقرر يُدرس بهذه اللغة، فيصبح الطالب قادرًا على استدعاء المعارف باللغة الأم(العربية)، بينما لا يستطيع ذلك باستخدام اللغة الإنجليزية، مما يؤثر أيضًا في درجة تقييمه. ولا تقيس الاختبارات التحصيلية التحريرية مهارات بعينها كالتواصل الشفوي والمناقشة وهي مهارات يمكن قياسها من خلال الاختبارات الشفوية كأحد أنواع الاختبارات التحصيلية.

مخرجات التعلم التي تقيسها الاختبارات الشفوية

- ❖ قياس مهارات يصعب قياسها بالاختبارات التحريرية، مثل: قدرة الطالب على المناقشة، ومهارات التواصل اللغوية، والمهارات التعبيرية، وتمثيل الأدوار.
- ❖ التحقق من إنجاز الطلاب لمخرجات التعلم غير المعرفية، مثل: مهارات التواصل اللفظي وغير اللفظي، والتعبير عن الأفكار والمشاعر في سياقات مألوفة للآخرين، والتي صنفها الإطار الوطني للمؤهلات ضمن المجال الثاني من مخرجات التعلم.
- ❖ تتناسب مع الطلاب ذوي الإعاقة والحالات الخاصة الناتجة عن إعاقات جسدية قد تمنعهم من الإجابة على الأسئلة التحريرية.

وتتسم الاختبارات الشفوية بميزات عديدة، كما أنها ذات محدودية في الاستخدام، وذلك على النحو التالي:

المميزات	محدودية الاستخدام
• قياس مدى عمق فهم الطالب للمعلومات، ومهارات التفكير العليا، وتقديم الأدلة على قيامه بهذه العمليات. • ملاحظة ردود أفعال متنوعة من قبل الطالب مثل تردده، أو ارتباكه في نطق كلمات معينة، والتي قد تؤخذ في الاعتبار عند تقييم درجة تمكن الطالب من مهارات تواصل مثل الإلقاء أو العرض. • تقديم التغذية الراجعة الفورية للطالب بناء على استجابته، مما يجعل الطالب قادرا على التعبير عن وجهة نظره والمناقشة، والتعبير عن أرائه بمزيد من الوضوح، وإضافة معلومات جديدة أو إعادة صياغة الإجابة.	• يستغرق تطبيقها وقتًا طويلًا، حيث تقييم الطلاب بصورة فردية، مما يؤدي إلى صعوبة استخدامها مع الأعداد الكبيرة للطلاب. • تباين صعوبة الأسئلة التي توجه للطلاب المختلفين، وصعوبة التحقق من تكافؤها. • عدم اتساق التقديرات وبخاصة في عدم وجود مصفوفة جيدة، ومحكات واضحة للتقدير.

المميزات	محدودية الاستخدام
• تحديد جوانب القوة والضعف والصعوبات في تحصيل الطلاب للتغلب عليها. • المرونة في قياس ومعاينة نطاق أكبر من مخرجات التعلم المراد قياسها. • تسمح للطالب بتعديل إجابته. • سهولة تطبيقها. • غير مكلفة ماديًا مقارنة بالاختبارات التحريرية.	• أثر الهالة للطالب ونظرة الممتحن لكل طالب في ضوء أدائه في المحاضرات يمكن أن ينعكس في تقديراته. • تأثر إجابات الطالب بخصائصه الشخصية كالخجل والتردد.

وعلى الرغم من عيوب الاختبارات الشفوية إلا أن الإعداد المسبق لها، وتقنين طرق إجرائها، وتحديد المدة الزمنية الفاعلة المستغرقة في التطبيق من شأنه أن يزيد من فاعليتها كأحد أدوات التقويم التي تمد أستاذ المقرر بمعلومات قيمة عن إنجاز الطالب لمخرجات التعلم غير المعرفية.

معايير إعداد وتوجيه أسئلة الاختبارات الشفوية

- ❖ مواءمة اختبار مع مخرجات التعلم المناسبة له، مثل مهارات التواصل الشفهي.
- ❖ الإعداد المسبق للاختبار من حيث كتابة الأسئلة، ووضوح الهدف منها، وتحديد العدد المناسب منها لقياس مخرجات التعلم.
- ❖ بناء مصفوفات التقدير/ التصحيح المناسبة في ضوء محكات الأداء، ومستوياتها للتحقق من استيفاء الطالب لأداء المهام، أو تحقيق المهارات المطلوبة وفقًا لمعايير محددة مسبقًا.
- ❖ تجنب أثر الهالة أو الحكم على الطالب من مظهره، أو سماته الشخصية، وليس أدائه.
- ❖ اقتصار الموقف الاختباري على قياس إنجاز الطالب لمخرجات التعلم، وتجنب تحويله إلى فرصة للجدال مع الطالب.
- ❖ توفير مناخ آمن عند توجيه السؤال للطالب.
- ❖ تقديم الأسئلة بطريقة منطقية ويفضل من العام إلى الأكثر نوعية وتخصصية.
- ❖ إعطاء وقت للطالب للتفكير قبل طلب الإجابة، وإخبار الطالب بالوقت اللازم للإجابة على السؤال، مثال: "في خلال ثلاث دقائق أشرح الفرق بين القياس والتقييم والتقويم"
- ❖ عدم مقاطعة الطالب أثناء الإجابة إلا إذا ابتعد عن موضوع السؤال، أو كانت المقاطعة لتوضيح الإجابة (ماذا تقصد، هل من أسباب أخرى.....).
- ❖ عدم التعليق على إجابة الطالب بكلمات محبطة (خطأ، كفى، .....).
- ❖ الانتقال إلى السؤال التالي في حال فشل الطالب في الإجابة على السؤال المقدم له، دون التركيز طويلًا على هذا السؤال، مع مراعاة إعطاء الوقت الكافي للطالب للتفكير في إجابة السؤال.

- تقدير درجات الاختبارات الشفوية.
- توفير الإمكانات والبيئة المناسبة لإجراء الاختبارات الشفوية التي تسهل إمكانية تسجيل الاستجابات، لضمان إعادة تقدير الدرجات، وحساب اتساق التقدير interrater reliability، ويمكن أيضًا تدوين الملاحظات أو التقديرات عن أداء الطلاب أثناء إجراء الاختبار أو بعده مباشرة.

أسئلة شائعة

❖ ما مدى ارتباط الاختبار التحريري بمجالات مخرجات التعلم طبقًا للإطار الوطني للمؤهلات؟

يقيس الاختبار التحريري مخرجات التعلم التي تنتمي لكل من مجال المعرفة والفهم والمهارات الإدراكية.

❖ كيف يتم التحقق من جودة الورقة الاختبارية؟

يتم ذلك بالتحقق من خلال:

- ✓ بناء الورقة الاختبارية وفق جدول المواصفات، والتأكد من مواءمة المفردات الاختبارية مع مخرجات التعلم التي تقيسها، والأوزان النسبية لها، والدرجات المخصصة لكل سؤال.
- ✓ مراجعة الورقة الاختبارية من قبل المراجع الداخلي، والتثبت من دقة المواءمة مع جدول مواصفات الاختبار المُعد مسبقًا (العمود الأخير في نموذج (٢- أ)، حيث يتم التحقق من توافق كل مفردة مع مخرج التعلم المستهدف بالقياس، وتعد هذه المواءمة مؤشراً لصدق الاختبار.
- ✓ مراجعة الورقة الاختبارية من قبل المراجع الخارجي بدقة، وذلك باستخدام نموذج٤: قائمة التحقق من جودة الورقة الاختبارية، ثم اعتمادها.

❖ كيف يساعد جدول المواصفات في إعداد الصور المتكافئة للاختبار؟

يمكن صياغة صورتين متكافئتين للاختبار استنادًا إلى بنود جدول المواصفات بحيث تتكافئين في:

- ✓ مخرجات التعلم
- ✓ المحتوى المتضمن
- ✓ مستوى صعوبة الأسئلة
- ✓ نمط السؤال المستخدم
- ✓ الدرجات المخصصة للاختبار
- ✓ زمن الاختبار.

❖ هل ينبغي إطلاع الطالب على جدول المواصفات للاختبارات؟

يمكن إطلاع الطالب فقط على الأوزان النسبية لمخرجات التعلم، والدرجات المخصصة لها، حيث أن هذا الإجراء يضمن شفافية التقويم من ناحية، وزيادة دافعية الطالب، وتوجيه تعلمهم وطرق استذكارهم، لتكون موجهة بمخرجات التعلم فيما يعرف بال outcome-based education، ولا ينبغي مشاركة أو اطلاع الطلاب على الموضوعات التي سيتم قياسها.

❖ ما هي إجراءات ضمان صدق الاختبار؟

- ✓ دقة المواءمة البنائية بين مخرجات التعلم والمفردات الاختبارية التي تقيسها وطرق التدريس، حيث إن قياس مخرجات التعلم تُعني قياس ما قام الطلاب باكتسابه وتحقيقه من خلال أنشطة التدريس والتعلم أثناء الفصل الدراسي، وعليه فإن القصور في استخدام استراتيجيات التدريس أو في تدريس المحتوى الذي يحقق مخرجات التعلم يؤثر سلباً في صدق الاختبار، وقياس المخرجات التي وضع لقياسها.
- ✓ دقة مواءمة المفردات الاختبارية وفقًا لجدول مواصفات الاختبار.
- ✓ التحقق من جودة الورقة الاختبارية، من حيث جودة صياغة الأسئلة، ووضوح التعليمات، وخلوها من أي غموض قد يتسبب في حدوث خلط في المفاهيم لدى الممتحن.
- ✓ تناسب طول الاختبار مع زمن الإجابة عليه، لأن الاختبارات الطويلة تحول الاختبار من اختبار قوة (يميز بين الطلاب في مدى إتقانهم للمحتوى ومدى تحقيقهم لمخرجات التعلم) إلى اختبار سرعة يقيس الفروق بينهم في سرعة الأداء، وعليه قد يلجأ الطالب إلى تخمين الاستجابة في حال كان الوقت غير كاف، مما يؤثر في صدق دلالة الدرجات.

❖ ما هي الاحتياطات التي يجب أن نأخذها لضمان ارتفاع نسبة ثبات الاختبار؟

أسباب انخفاض الثبات	الاحتياطات
غموض صياغة أسئلة الاختبار	مراجعة الاختبار وتعديله في ضوء التغذية الراجعة من المراجع الداخلي.
تطرف صعوبة الأسئلة	اختيار مفردات متوسطة الصعوبة تتناسب مع مخرجات التعلم المستهدفة طبقًا لجدول المواصفات.
قلة عدد الأسئلة	زيادة طول الاختبار بزيادة عدد الأسئلة وتضمين الأسئلة الموضوعية، فالاقتصار على الأسئلة المقالية قليلة العدد يقلل من قيمة معامل ثبات الاختبار، شريطة أن يتناسب الوقت المخصص للاختبار مع عدد الأسئلة.
استخدام أسئلة اختيارية، أي أن نقدم للطالب سؤالين ونطلب منه أن يختار أحدهما ليجيب عليه	توحيد الأسئلة المطلوب الإجابة عليها بالنسبة لكل الطلاب؛ لأن ذلك يساعد على كشف الفروق الفردية بينهم وارتفاع نسبة ثبات الاختبار.
عدم ملاءمة الموقف الاختباري يزيد من تأثير الأخطاء العشوائية في القياس.	ضبط الموقف الاختباري ووضع خطة طوارئ. تقليل عوامل الضوضاء والتشتيت في قاعة الاختبار، مثل: التعليمات الشفوية أثناء الاختبار، والسماح بدخول الطلاب المتأخرين، وعدم مناسبة الإضاءة أو درجة الحرارة.
تجانس الطلاب في المستوى	العمل على زيادة الأسئلة المميزة في الاختبار للكشف عن الفروق بين الطلاب، حيث إن معامل الثبات يرتبط طردياً بمقدار التباين في درجات الطلاب.

❖ لماذا الاختبارات المقالية أقل ثباتًا من الاختبارات الموضوعية؟

يرجع السبب في كون الاختبارات المقالية أقل ثباتًا من الاختبارات الموضوعية إلى الذاتية في التقدير أو التصحيح أو قلة عدد الأسئلة في الاختبار.

❖ ما هي إجراءات ضمان عدالة الاختبار؟

- ✓ التحقق من تكافؤ الاختبارات البديلة مع الاختبارات الأساسية في مستوى الصعوبة، وذلك باستخدام جدول المواصفات، بحيث يقيس كل منهما نفس مخرجات التعلم، وبنفس شروط تحقق التكافؤ التي سبق الإشارة إليها.
- ✓ تجنب منح بعض الطلاب فرص إضافية لأداء الاختبار دون غيرهم، فيجب إعلان معايير أداء الاختبارات والخاصة بالحرمان والاعتذار لجميع الطلاب، وتطبيقها وفقًا لما ورد في دليل السياسات.
- ✓ تجنب تقديم أي معلومة عن الاختبار مثل الأوزان النسبية لمخرجات التعلم، أو نوع المفردات بشكل فردي للطلاب خلال الساعات المكتبية، فإذا تقرر إعلان هذه المعلومات فيجب أن يتم ذلك من خلال المحاضرات أو نظام إدارة التعلم، لضمان تكافؤ الفرص.
- ✓ استخدام أمثلة واضحة وشائعة بين الجميع، وتجنب استخدام أمثلة متحيزة لبيئة معينة، يفهمها سكان هذه البيئة دون غيرهم.
- ✓ مواءمة الاختبارات للطلاب ذوي الإعاقة والحالات الخاصة.

المراجع

- أبو فودة، باسل، نجاتي بني يونس (٢٠١٢). الاختبارات التحصيلية: مفهومها، كيفية إعدادها، أسس بنائها وتكوينها، وتطبيقات ميدانية. عمان، الأردن: دار المسيرة.
- العزاوي، رحيم (٢٠٠٩). القياس والتقويم في العملية التعليمية. عمان، الأردن: دار دجلة ناشرون وموزعون.
- المحاسنة، إبراهيم، مهيدات، عبد الحكيم (٢٠٠٩). القياس والتقويم الصفي، عمان، الأردن: دار جرير للنشر والطباعة.
- جامعة تبوك. دليل التقويم الواقعي في المقررات الجامعية. وحدة القياس والتقويم.
- سبع ابو لبدة (٢٠٠٨). مبادئ القياس النفسي والتقييم التربوي. عمان، الأردن: دار الفكر.
- سيد، سعاد (٢٠١٧). التقويم التربوي، رؤى وأساليب. القاهرة، مصر: دار الكتاب الحديث.
- عبد الحميد، جابر (١٩٩٦). التقويم التربوي والقياس النفسي. القاهرة، مصر، دار النهضة العربية.
- عثمان، السعيد (٢٠١٤). الاختبارات التحصيلية كأداة للتقويم. القاهرة، مصر: مؤسسة طيبة للنشر والتوزيع
- علام، صلاح (٢٠٠٤). التقويم التربوي البديل. أسسه النظرية والمنهجية وتطبيقاته الميدانية. القاهرة، مصر: دار الفكر العربي.
- علام، صلاح (٢٠٠٥). القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة، مصر: دار الفكر العربي.
- علام، صلاح الدين محمود. (٢٠١١). القياس والتقويم التربوي والنفسي، أساسياته وتطبيقاته وتوجهاته المعاصرة، ط٥، القاهرة: دار الفكر العربي.
- علام، صلاح (٢٠١٥). القياس والتقويم التربوي في العملية التدريسية. عمان، الأردن: دار المسيرة للنشر والتوزيع.
- عودة، أحمد سليمان. (٢٠١٠). القياس والتقويم في العملية التدريسية. الأردن: دار إربد، ط٤.
- ماهر اسماعيل، محب محمود (٢٠٠٣). التقويم التربوي: أسسه وإجراءاته. الرياض، السعودية: مكتبة الرشد.
- ملحم، سامي (٢٠٠٠). القياس والتقويم فى التربية وعلم النفس. عمان، الأردن: دار المسيرة للنشر والتوزيع.

Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for assessment also good for learning? *Journal of Applied Research in Memory and Cognition*, *7*(3), 323-331.

Flateby, T. L. (2014). A guide for writing and improving achievement tests. *Office of the Provost, University of South Florida, Tampa, FL Retrieved*.

Gronlund. N. (2006). Assessment of Student Achievement. United States: Pearson Education Inc.

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: principles, policy & practice*, *10*(2), 169-207.

Kennedy, D. (2006). Writing and using learning outcomes: a practical guide. University College Cork.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and social psychology review*, *15*(1), 28-50.

Miller, M., Linn, R. & Gronlund, N. (2013). Measurement and Assessment in Teaching. United States: Pearson Education Inc.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist*, *37*(2), 91-105.

University of Washington. Office of Educational Assessment. Retrieved from: http://www.washington.edu/assessment/scanning-scoring/scoring/reports/item-analysis/

Young, J., So, Y., & Ockey, G. (2013). Guidelines for best test development practices to ensure validity and fairness for international English language proficiency assessments. *Educational Testing Service*.

Wang, Z., Gong, S. Y., Xu, S., & Hu, X. E. (2019). Elaborated feedback and learning: Examining cognitive and motivational influences. *Computers & Education*, *136*, 130-140.

William, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in education: principles, policy & practice*, *11*(1), 49-65.

مركز جودة التقويم والاختبارات

Center For Quality of Assessment and Examinations