مفهوم الصدق في الاختبارات التحصيلية: الخاصية أم المشكل؟

د. ربيعة جعفور
جامعة قاصدي مرباح ورقلة (الجزائر)

الملخص:

يتناول هذا المقال الوقوف على مفهوم الصدق وحقيقة افتقاد الكثير من الاختبارات التحصيلية لتحقيق هذا المؤشر مما يدلّل على وجود مشكل لدى مصمّمي الاختبار في الوقوف على أهم أسس بناء الاختبار التحصيلي الجيد أو حتى تبني النظرية الأحـــدث في القياس ليتمكن من تقويم مخرجات العملية التعليمية ويحقق الشمولية والموضوعية في الأحكام والقرارات الصادرة عـــن هـــذه العملية برمتها.

الكلمات المفتاحية: الاختبارات التحصيلية، الصدق.

Abstract:

This article examines the stand on the concept of validity and the fact that the lack of a lot of achievement tests to achieve this indicator as evidence of a problem with the designers of the test stand on the most important foundations of the achievement test good or even adopt a theory the latest in measurement to be able to evaluate the outputs of the educational process and achieve universality and objectivity in provisions and decisions issued by this whole process.

The Key words: Achievement tests**,** validity**.**

Résumé:

Cet article examine la position sur la notion d'honnêteté et le fait que l'absence d'un grand nombre de tests de rendement pour atteindre cet indicateur, ce qui démontre qu'il ya un problème avec les concepteurs du test se tiennent sur les fondements les plus importants de la bonne réalisation du test de rendement ou même adopter une théorie plus tard en mesure d'être en mesure d'évaluer les résultats du processus d'enseignement et d'atteindre l'universalité et l'objectivité dans jugements et décisions émises par l'ensemble de ce processus.

Les Mots clés: tests de rendement, la validité

مقدمة:

تعد الاختبارات التحصيلية الوسيلة الأكثر تداولا في مؤسساتنا التعليمية والتربوية ذلك أنها تساعد على إصدار أحكام يتم على أساسها انتقال التلميذ من مستوى لآخر أو إنهاؤه لمرحلة تعليمية، ولكن الملاحظ للواقع التعليمي ونظام الانتقال في بيئتنا العربية خاصة يجد أحيانا سوء فهم واستخدام لوسيلة الامتحان مما يترتب عليه آثارا بعيدة المدى تضر بالتلميذ كما الأسرة والمجتمع على حد سواء، فلقد أصبحت الامتحانات الفصلية أو امتحانات نهاية العام الدراسي حسب نزال كما أورد في(فؤاد علي العاجز، 2008، ص13) هي التي توجه التعليم وهي التي تدعو التلاميذ والمدرسين إلى الاهتمام بجميع ما تتطلبه، وإهمال جميع ما لا يدخل في نطاقها، ولا تقوى على قياسه، وبذلك صارت الامتحانات غاية في حد ذاتها بدلا من أن تكون وسيلة لتحقيق أهداف التربية الشاملة؛ كما أن المبالغة في استخدامها وسوء فهم الهدف منها أفقدت عملية التعليم والتعلم كثيرا من قيمتها في إعداد مواطنين صالحين؛ مما نتج عنه بالمحصلة كثير من المظاهر السلبية كاقتصار التدريس على جانب التلقين وشيوع الملخصات والكتب الخارجية والدروس الخصوصية التي تساعد التلاميذ على اجتياز الامتحان، بينما أهمل ما له علاقة بالتفكير الناقد وغيره من المستويات المعرفية العليا؛ لذا فقد وجهت لها انتقادات كثيرة ذكرها (فؤاد علي العاجز2008، محمد حسين سعيد حسين2005، محمد أحمد محمد إبراهيم غنيم2003) تتعلق بمدى كفاءتها من حيث الثبات والموضوعية والصدق وتركز في معظمها على المستويات المعرفية الدنيا وتهمل المستويات المعرفية العليا، أضف إليه ضعف القيمة التشخيصية للامتحانات وما تسببه من ارتفاع مستوى القلق والخوف لدى التلاميذ وأوليائهم والذي قد يصل حد الانهيار العصبي.

يوصي كل من (Hambleton& Wedman, 1997) كما أورد في (محمد أحمد محمد إبراهيم غنيم2003) بضرورة التخلص من المشكلات المرتبطة بالاختبارات المعتادة والتي تتمثل في عدم موضوعيتها وعدم قدرتها على تقدير النطاق الواسع للمحتوى، كما أنها لا توفر تقييما تشخيصيا جيدا، وذلك من خلال تبني الاتجاهات الحديثة في القياس النفسي التي تمد الممارسين بأطر تسهم في تطوير عملية التقويم وتوفر موضوعية القياس وتركز على تقويم الكفايات بدلا من تقويم الأداء، والاستعانة بالتكنولوجيا الحديثة في تطوير النماذج السيكومترية الحديثة التي تسهم في تطوير طرق بناء الاختبارات وتحليل بنياتها، ويضيف كل من (Hambleton& Slater, 1997) أن الاختبارات في الوقت الحالي تمر بتغيرات أساسية وإحدى هذه التغيرات هو الانتقال من الاعتماد على القياس الكلاسيكي إلى تبني النماذج الحديثة للتغلب على عيوب الاختبارات التقليدية ومنها تباين درجات الأفراد بتباين مستويات صعوبة المفردات.

وعليه فان الفهم الصحيح للامتحانات من شأنه أن يحقق فوائد كثيرة ذلك أن نتائجها تعد معلومات فاصلة لاتخاذ القرارات التربوية المتصلة بتحسين العملية التعليمية، وتوجيهها بما يخدم أغراض المجتمع وغاياته، ومن هنا برزت الحاجة إلى استخدام أساليب ذات كفاءة سيكومترية قصد تحقيق التقييم الدقيق والشامل لأهداف العملية التعليمية والتربوية، فما هي أسس الاختبار الجيد؟ وإلى أي مدى يمكن توفر خاصية الصدق وفق النماذج المعتمدة في تصميم الاختبارات التحصيلية وكيف يمكن تفسير درجاتها بالشكل المناسب؟ وما هي الأطر النظرية الحديثة التي تساعد على الفهم الصحيح للصدق ومؤشراته؟

أولا- مفهوم الاختبارات التحصيلية:

تعتبر الاختبارات التحصيلية أداة أو وسيلة مقصودة منها تقويم مخرجات العملية التعليمية فهي " إجراء منظم لتحديد مقدار ما تعلمه التلاميذ في موضوع ما في ضوء الأهداف المحددة ويمكن الاستفادة منه في تحسين أساليب التعلم ويسهم في إجادة التخطيط وضبط التنفيذ وتقويم الانجاز" (عماد عبد الرحيم الزغلول، شاكر عقلة المحاميد، 2007، ص172)

ومن ناحية أخرى فالاختبارات مرتبطة تماما بمفهوم النجاح والفشل فهي" اختبارات مصمّمة للكشف عن درجة نجاح التلميذ في مادة أو مواد كان قد تعلّمها مسبقاً، وتتضمن هذه الاختبارات عادة على أسئلة تكشف عن المهارات والمعارف الأساسية للمادة أو المواد"(نبيل عبد الهادي، 2001، ص472).

وعليه يمكن تحديد الهدف من الاختبار التحصيلي في عدد من العناصر منها النقل والترفيع للمستوى الأعلى، والتسكين والتشخيص وتحديد المستوى، والاختيار، والانتقاء الدراسي والمهني وتحديد مستوى التحصيل (يوسف قطامي وآخرون، 2000، ص 885)

وهذه التعاريف لا تعن أن هناك شكلا وحيدا يجسد هذا المفهوم بل تتنوع الاختبارات التحصيلية تبعا لعدة اعتبارات.

ثانيا- تصنيف الاختبارات التحصيلية:

يوجد العديد من تصنيفات الاختبارات التحصيلية وذلك حسب درجة التقنين، عامل السرعة، الأسلوب المستخدم، تفسير الدرجة، فهي من حيث درجة التقنين (اختبارات مقننة، اختبارات غير مقننة) ومن ناحية عامل الوقت أو السرعة (اختبارات سرعة، اختبارات دقة) وعلى أساس الأسلوب المستخدم (اختبارات شفوية، اختبارات أدائية) أما على أساس تفسير الدرجة على الاختبار فهي (اختبارات معيارية المرجع، اختبارات محكية المرجع) ولعل هذا التصنيف الأخير للاختبارات التحصيلية هو الأكثر شيوعا في تقدير نتاجات العملية التعليمية.

1. الاختبارات معيارية المرجع:

تندرج الاختبارات معيارية المرجع على ما يعرف بالتقويم معياري المرجع وهو يعرف بأنه" إصدار حكم على مستوى أداء المتعلمين أو المتدربين ومدى إنجازهم وفاعلية تعلمهم باعتبار أداء مجموعة التلاميذ أو المتدربين الممثلين لهم في الظروف والخلفيات الاجتماعية والاقتصادية، والديمغرافية، والعمرية بهدف المقارنة بين أداء كل منهم إلى مجموعته المرجعية"(يوسف قطامي وآخرون، 2000، ص920)

يقوم هذا النوع من التقويم على مجموعة من الأسس تستند إلى نظرية القياس الكمية وهو يعتمد على المعالجة الإحصائية لتفسير الدرجة الخام التي يحصل عليها التلميذ وهذه الأسس هي حسب (يوسف قطامي وآخرون،2000، ص920-921):

- التوزيع الطبيعي للظاهرة التحصيلية أي 68% في الوسط و16% على الطرفين
- يقارن أداء التلميذ في الاختبار بأداء زملائه أو الفئة التي يندرج ضمنها أو الفئة التي تفوقه في الأداء أو تقل عنه.
- يفرز نسبة الناجحين ونسبة الراسبين.
- تفسر علامات التلميذ التحصيلية على وفق متوسط المجموعة التي يأتي منها التلميذ
- المتوسط يشكل معيارا على أساسه يمكن الحكم على أداء التلميذ بعد تحويل علامته الخام إلى علامة معيارية
- يقيس هذا الاختبار أهدافا نهائية وشاملة
- يوصف الاختبار لهذا الغرض بأنه اختبار معياري

يتبين إذن أن الاختبارات معيارية المرجع تعتمد على مقارنة أداء التلميذ بأداء مجموعته الصّفية، وبالتالي فإنه من المحتمل أن يختلف موقع التلميذ باختلاف خصائص مجموعته المعيارية، بالإضافة إلى أنه ليس بالضرورة أن يمثّل متوسط المجموعة المعيارية الأداء المطلوب للنجاح، وتهتم هذه الاختبارات أيضاً بإبراز الفروق الفردية بين التلاميذ دون النظر إلى مدى إتقان التلميذ للمهارات والمعلومات المراد قياسها وذلك لتأهيله لبرامج تدريبية وتعليمية جديدة،

وبالتالي فهي لا تساعد على تقويم وتشخيص كل من كفاية التلميذ وبعض البرامج التدريبية، أي أنها لا تزود المعلم بمعلومات دقيقة تساعد على اتخاذ القرارات التربوية المناسبة حول مستوى كفاية التلميذ والمنهج الدراسي؛ لعبت هذه الانتقادات دوراً في ظهور تطورات جديدة في منهجية الاختبارات والمقاييس التربوية أدت إلى ظهور اتجاه آخر أطلق عليه القياس محكي المرجع(Criterion-Referenced-Measurement) حيث سجّلت أول مقالة لجليزر في عام(1963 م) تناقش هذا الاتجاه(خالد بن حسن شيبان التميمي، 1999، ص2-3).

2. الاختبارات محكية المرجع:

تندرج الاختبارات محكية المرجع تحت التقويم المحكي المرجع الذي يعرف بأنه" إصدار حكم أو قيمة على مستوى إنجاز وتحقيق المتعلمين أو المتدربين للنواتج المرصودة" (يوسف قطامي وآخرون،2000، ص924) وسمي هذا التقويم بمحكي المرجع لأن المدرس أو المدرب يتبنى الأهداف المرصودة بمثابة محكّات تحدد مدى فعالية التعلم والتحسن في المادة أو المساق أو البرنامج التدريبي وهو يستند إلى مجموعة الافتراضات التالية حسب (يوسف قطامي وآخرون،2000، ص924):

- يستطيع أيّ متعلم أن يحقق الأهداف التعليمية إذا توافرت لديه الدافعية للتعلم والاستعداد لذلك
- إن المتعلم الذي يعي الأهداف التعليمية يستطيع أن يقوم بالنشاطات اللازمة لتحقيقها وذلك بسبب صياغة الأهداف على وفق أسس منطقية تتعلق بمستوى تدرج المواد في المادة على أسس نمائية تحدد مستوى تفاعلها مع الخبرات والمواد التي تقدم له.
- إن الأهداف بما تتضمنه من تخطيط دقيق للنواتج وللزمن المستغرق تساعد المتعلم على الوصول إلى أداء يعكس إمكاناته وقدراته التحصيلية.
- إن توفير الأسلوب التدريسي المناسب والأهداف المرصودة والمحددة بدقة والزمن المناسب لتعلم التلاميذ يبرر الاستناد على الأهداف كمحك للحكم على أداء التلاميذ.
- إن صياغة الأهداف السلوكية صياغة سلوكية دقيقة قابلة للملاحظة والقياس تساعد على توفير أساس لاعتمادها كمحك يقيم عليها أداء التلاميذ.

يهدف القياس محكي المرجع إذن إلى تصنيف التلاميذ إلى مجموعتين الأولى متقنة (Mastery) والثانية غير متقنة(Non-Mastery) للمهارات والمعارف وذلك اعتماداً على (مستوى أداء – معيار – محك – درجة قطع – درجة نجاح) محددة تحديداً مسبقاً، ويتمثل هذا المحك(Criterion) في نسبة عدد الأسئلة التي يجب أن يقوم التلميذ المتقن بالإجابة عليها بصورة صحيحة، وبالتالي إذا كانت درجة التلميذ في الاختبار تزيد أو تساوي درجة القطع off (Cut Score) اعتبر التلميذ متقناً للمهارات وإذا لم تكن كذلك اعتبر التلميذ غير متقن(خالد بن حسن شيبان التميمي، 1999، ص3)

وعليه تختلف الاختبارات القائمة على أساس التقويم المحكي المرجع عن نظيرتها القائمة على التقويم المعياري المرجع في كون الاختبار مرجعي المعيار هو اختبار بعدي ونسبي؛ أي يظهر بعد الأداء ونسبي لأنه يتأثر بمتوسط العينة، أما الاختبارات مرجعية المحك فهي قبلية ومطلقة ذلك أن نقطة الفصل في تحديد مقدار القبول هي التخطيط المبكر لما يجب ألا ينخفض عنه الأداء، وعلى ذلك فإن قدرة جميع الأفراد أو نسبة منهم على بلوغ هذا القدر أو تجاوزه أو عدم بلوغ أي منهم له لا يدخل كعامل محدد لمقدار الأداء المطلوب بوصفه أداء مناسبا، ويعني هذا أننا أمام مفهوم مطلق لا يعتمد على عوامل متغيرة تتعلق بمستوى الصعوبة أو مستوى الأداء أو طبيعة العينة؛ كما أن المتوسط في الاختبارات مرجعية المعيار عبارة عن نقطة متغيرة، كما أنه في الاختبارات مرجعية المحك قد يترتب على المعيار المسبق أن ينجح أو يفشل كل المفحوصين في الاستجابة على الاختبار(صفوت فرج، 2000، ص210-211).

تجدر الإشارة إلى أن كثيراً من علماء القياس والتقويم التربوي تناولوا أوجه الشبه والاختلاف بين هذه الاتجاهات المختلفة في القياس المعاصر وذلك لأهمية هذه الاتجاهات في قياس وتقويم مخرجات العملية التعليمية، كما أشار كل من جليزر، وإيبل، وبابام وهيوسك، وهامبلتون وآخرون، Hambleton et al, 1978 Papham&Husek, 1969; Ebel, 1979; Glaser, 1994) إلى وجود عدد من الاختلافات بين الاختبارات معيارية المرجع والاختبارات محكية المرجع، والتي يمكن تلخيصها في العناصر التالية (كما ورد في: خالد بن حسن شيبان التميمي، 1999، ص14- 15):

1.الهدف الرئيس للاختبار:

يهدف الاختبار محكي المرجع إلى مقارنة أداء التلميذ بمحك محدد مسبقاً وذلك لمعرفة الأهداف التي حققها والتي لم يحققها التلميذ، أما الاختبار معياري المرجع فيهدف إلى مقارنة أداء التلميذ بأداء مجموعته المعيارية من خلال تحديد عدد الأسئلة التي أجاب عليها التلميذ بشكل صحيح.

2.استخدام الاختبار:

يستخدم الاختبار محكي المرجع للتأكد من تحقيق التلميذ للأهداف السلوكية المحددة، أما الاختبار معياري المرجع فيعتمد على تزويدنا بمؤشرات عن درجة نجاح التلميذ، وبالتالي إظهار الفروق الفردية بين التلاميذ.

3. خصائص أسئلة الاختبار:

تتجمع الأسئلة في الاختبار محكي المرجع حول عدد محدد من الأهداف، بينما تنتشر الأسئلة في الاختبار معياري المرجع بشكل واسع حول نطاق الأهداف.

4.تفسير الأداء:

تفسر الدرجة في الاختبار محكي المرجع بناء على درجة القطع(الإتقان)، بينما تفسر درجة التلميذ في الاختبار معياري المرجع بناء على درجات معيارية تمثل موقع التلميذ بين مجموعته الصفية والرتب المئينية.

5.بناء الفقرات:

يعتمد بناء الفقرات في الاختبار محكي المرجع على مدى تحقيق المفردة للهدف السلوكي التعليمي، بينما يعتمد بناؤها في الاختبار معياري المرجع على تباين الدرجات ومعامل التمييز، ويتجنب اختيار المفردات الصعبة جداً والسهلة جداً.

6.التقويم:

يسعى الاختبار محكي المرجع إلى تحديد جوانب الضعف والقوة في كل من أداء التلاميذ والبرامج التعليمية، بينما يسعى الاختبار معياري المرجع إلى تقويم أداء التلميذ ومقارنته مع أ داء التلاميذ الآخرين.

مما يعن أنه يتوجب توفر مجموعة من الأسس تضمن فعالية الاختبار التحصيلي بغض النظر عن الاتجاه الذي يتبناه مصمم الاختبار.

ثالثا- أسس الاختبار التحصيلي الجيد:

إن تحليل نظرية القياس المتضمنة للفروق الفردية في الأداء إزاء أي خبرة أو موقف تعليمي أو تدريبي يمكن أن يحددا عددا من الأسس وهي حسب (يوسف قطامي وآخرون،2000، ص884):

- يشتمل الاختبار على عينة ممثلة من الأسئلة تقيس الأهداف والمحتوى حسب الأهمية والوزن.
- يصمم الاختبار ليقيس بوضوح النتاجات التعليمية المشتقة من أهداف المقرر.
- تحديد نوع فقرات الاختبار على وفق المحتوى والأهداف.

- تستثمر نتائج الاختبار في مراقبة تعلم التلاميذ وتحسينه وتطويره.
- الاختبار الأكثر ملائمة هو الاختبار الذي تتوافر فيه خصائص الاختبار الجيد.
- تزود نتائج الاختبارات بتغذية راجعة تصحيحية وتعزيزية.
- يزود الاختبار بسبل النمو المهني للمدرس.
- تفسر نتائج الاختبار بحذر ودقة.

وقد تم استنتاج هذه الأسس على وفق طبيعة المتعلم والخبرة ودور المدرس وأهداف الدراسة، وفلسفة الدراسة والممارسات التدريسية التي يؤديها المدرس بهدف تحقيق الأهداف وإحداث التغير المناسب لدى التلاميذ، ولنفترض أن كل مدرس يقوم ببناء اختباره على وفق أسس خاصة يضعها فهل سنحصل على نتائج متشابهة وبمعنى آخر هل سيقيس الاختباران نفس الشيء مما يطرح مفهوم صدق الاختبار.

رابعا- مفهوم الصدق كخاصية للاختبار التحصيلي:

هناك شروط ومواصفات يجب أن تتوفر في أداة القياس لكي تعتبر أداة معتمدة وتؤدي إلى نتائج صادقة حيث أن صدق النتائج يساعد المسؤول على اتخاذ قرارات ثابتة، وموضوعية، وعادلة، وحيث يعد الاختبار أكثر الأدوات شيوعًا في المدارس وسواء كان هذا الاختبار من إعداد المعلم، أم من إعداد الخبراء والمختصين كالاختبار المقنن، فيجب أن تتوفر فيه شروط ومواصفات الاختبار الجيد وأهمها الصدق.

يعد الصدق إذن محور عملية القياس إذ أنه يشير إلى أن الاختبار يقيس ما وضع لقياسه وأن إجراءات التقويم المسؤولة يجب أن تكون صادقة، وصدق الاختبار يتعلق عادة بمدى خدمته للأغراض التي وضع من أجلها، كما يقصد بصدق الاختبار مدى صحة التفسيرات المبنية على نتائج الاختبار؛ حيث يرى(كرونباخ Cronbach1960) أنه بقدر اكتمال تفسير درجة الاختبار للسمة المعنية والثقة في هذا التفسير بقدر صدق الاختبار، وهو يربط بذلك بين الدرجة على الاختبار وقدرتها التفسيرية؛ ومن هنا فإننا نستطيع القول أن الصدق من حيث هو سمة أو خاصية من خصائص الاختبار لا يرتبط بالاختبار نفسه بل يعتمد على الغرض الأساسي الذي يستخدم الاختبار من أجله وبالقرار الذي سوف يتخذ بناء على نتائج هذا الاختبار، ويذكر الطريري (1997) في هذا الصدد تعريفا للصدق حسب ما ورد في معايير (1985) لـ AERA و APA أن الاختبار صادق بالدرجة التي تكون الاستنتاجات المبنية عليه مناسبة وذات دلالة وفائدة؛ وهذا يعني أنه عندما نقول صدق الاختبار فإننا نقصد صدق تفسير الدرجة لمستوى الخاصية أو السمة أو القدرة المراد قياسها، فالصدق يتعلق إذا بمدى فائدة أداة القياس في اتخاذ قرارات تتعلق بغرض أو أغراض معينة(سعيد حسن آل عبد الفتاح الغامدي، 2003، ص 27) ولما كانت المدرسة الحديثة تهدف إلى تقييم سلوك المتعلم من عدة نواحي، فإن الاختبارات الصادقة هي التي تحقق عدة غايات أهمها حسب (رجاء محمود أبو علام، 1978):

1. تشخيص الصعوبات في عملية التعليم
2. تقرير مدى تحقيق التعلم الأهداف التعليمية المقبلة.
3. التنبؤ بنجاح المتعلم في النشاطات التعليمية المقبلة.
4. اتخاذ قرارات إدارية وتربوية مختلفة.

وبهذا فالصدق يعتبر من أهم خصائص الاختبار الجيد على الإطلاق مما يعني أنه لابد من مراعاة بعض المحاذير عند استخدام مفهوم الصدق والتي يعدها البعض خصائص للصدق وهي حسب (رجاء محمود أبو علام 2004، عفاف بنت راضي مشخص اللحياني، 1429- 1430):

1. **أن الصدق يتوقف على عاملين** هما:

أ. الغرض من الاختبار، أو الوظيفة التي ينبغي أن يقوم بها

ب. الفئة أو الجماعة التي سيطبق عليها الاختبار.

2. **الصدق صفة نوعية:** أي خاصة باستعمال معين(بالغرض الذي من أجله وضع الاختبار) إذ أن خاصيته ليست ذات صفر مطلق، وعليه يكون اختبار التحصيل في مادة ما صادقا، إذا كان يقيس تحصيل التلميذ في تلك المادة، بينما لو طبق الاختبار نفسه في مجال آخر (كالذكاء مثلا) فإن نتائجه سوف تختلف.
3. **الصدق صفه نسبية،** أو متدرجة، وليست مطلقة، فلا يوجد اختبار عديم الصدق، أو تام الصدق، بمعنى أن الصدق له درجه فلا يوجد امتحان صادق، وآخر غير صادق، بل هناك امتحان ذو درجة عالية من الصدق، وآخر ذو درجة متوسطة، وآخر ذو درجة متدنية.
4. **الصدق صفه تتعلق بنتائج الاختبار** وليس بالاختبار نفسه، ولكننا نربطها بالاختبار من قبيل الاختصار أو التسهيل.
5. **يتوقف صدق الاختبار على ثباته،** أي على إعطاء النتائج نفسها تقريبًا في كل مرة يطبق فيها على مجموعة بعينها، فإذا تغير معامل الثبات كان الاختبار غير صادق.

خامسا- أنواع الصدق وطرق تقديرها:

حددت لجنة مشتركة من الجمعية الأمريكية لعلم النفس APA والجمعية الأمريكية للبحث التربوي AERA والمجلس القومي الأمريكي للقياس التربويNCME في كتاب معايير الاختبارات النفسية والتربوية(1999) ثلاث طرق رئيسية للصدق هي(صدق المحتوى، الصدق المرتبط بالمحك، صدق التكوين الفرضي) وهي تناظر ثلاث أنواع من القرارات والأحكام الصادرة عن درجات أدوات القياس، وهذه القرارات هي حسب (عفاف بنت راضي مشخص اللحياني1429- 1430، سعيد حسن آل عبد الفتاح الغامدي 2003):

1. تحديد كيفية أداء المفحوص في الوقت الحاضر في نطاق شامل لمواقف سلوكية تمثلها فقرات المقياس تمثيلا جيدا.
2. التنبؤ بالأداء المستقبلي أو تقدير مكانة المفحوص في أحد المتغيرات ذات الأهمية.
3. الاستدلال على درجة تملك المفحوص لسمة معينة من حيث هي تكوين فرضي.

i.صدق المحتوى Content Validity:

يطلق عليه أحيانا اسم الصدق المنطقي، أو صدق عينة السلوك أو الصدق بحكم المفهوم وهو متعلق بالصفات الداخلية للاختبار، ويستخدم هذا النوع من الصدق عندما يكون الهدف من تطبيق الاختبار وصف وتحديد الوظيفة المراد تحقيقها من الاختبار، لذا فإن الصدق الوصفي يعتبر شرطاً لازماً لحساب صدق المحتوى، ويعتبر الهدف الأساسي الذي يدعو الباحث لاستخدام هذا النوع من الصدق هو معرفة ما إذا كانت:

1. الخاصية ممثلة في مجموعة بنود وبشكل مناسب.
2. البنود تمثل المجالات الفرعية للخاصية وأبعادها.
3. مدى كفاية الاختبار في تمثيل مجموعة من السلوكيات التي سيجري عليها التعميم بناء على نتائج ذلك الاختبار

كما يلجأ الباحث إلى هذا النوع من الصدق في حال عدم توفر أدوات تقيس الخاصية التي يريد قياسها، أي بمعنى في حال تصميم أداة جديدة؛ وغالبا ما يقدر صدق المحتوى باختبار الخطة والإجراءات المستخدمة في بناء الاختبار، هل الإجراء يتبع مفهوما منطقيا يؤكد المحتوى بشكل ملائم؟ ثم هل أن بناء الاختبار يؤكد أن تجميع البنود سوف يمثل مهارات ملائمة ومطلوبة؟ وغيرها من الأسئلة التي تطرح هنا مثل:

- هل هناك عرض واضح للمهارات التي يمثلها الاختبار؟
- ما هي الطريقة التي تم استخدامها وتنفيذها لتحديد المحتوى المرغوب تقييمه؟
- ممن تكونت مجموعة الخبراء الذين قاموا باختبار الصدق؟ وكيف كان التحكيم؟
- ما مدى التشابه بين هذا المحتوى والمحتوى الذي تهتم باختباره؟

وجدير بالذكر أن بناء الاختبارات التحصيلية يبدأ باختيار الفقرات الملائمة ويسبق ذلك بحث شامل ومتعمق ومنتظم لخطط المادة ومقرراتها بالإضافة إلى استشارة الخبراء المختصين، وبناء على المعلومات التي تم جمعها في هذه المرحلة تكون مواصفات الاختبار قد تم إعدادها، بحيث تحدد هذه المواصفات جوانب المحتوى أو المواضيع التي يجب أن يشملها الاختبار، وكذا الأهداف السلوكية أو العمليات التي سيتم قياسها والأهمية النسبية لكل من هذه المواضيع، والعمليات؛ وبالتالي تحديد عدد الفقرات لكل موضوع أو عملية (عفاف بنت راضي مشخص اللحياني، 1429- 1430، ص83).

ويؤكد عبد الرحمن(1997) (كما ورد في: سعيد حسن آل عبد الفتاح الغامدي،2001، ص 28) بأن محتوى المقياس يكون صادقا طالما أنه يشمل جميع عناصر القدرة المطلوب قياسها ويمثلها، ولهذا فإن تقدير صدق المحتوى يتطلب أدلة منطقية وليست إحصائية، حيث يتم التحليل المنطقي لمحتويات المقياس ومطابقتها مع محتويات الجانب السلوكي المقاس من قبل المحكمين ويفضل الاعتماد على أكثر من محكم للحصول على تقديرات متجانسة ومتسقة، حيث يقومون بتحديد مكونات المجال السلوكي سواء كانت معارف أو مهارات أو عمليات تحديدا واضحا، وتحديد مكونات المجال السلوكي الفرعية والأهمية النسبية لكل منهما ثم يقوم المحكم بتقدير وتقييم كل بند من بنود المقياس في ضوء مكونات المجال السلوكي؛ ويستند تقدير صدق المحتوى إلى ثلاث فروض صاغها ليون(Lenon) (أورد في: عفاف بنت راضي مشخص اللحياني، 1429- 1430، ص84) كالتالي:

1. ينبغي أن يكون المجال الذي يختبر فيه المفحوص محددا ويمكن تعريفه تعريفا دقيقا
2. يمكن اختيار عينة من الأسئلة من هذا المجال بطريقة هادفة ومناسبة
3. يمكن تحديد عينة الأسئلة وطريقة المعاينات المستخدمة بدقة كافية لكي يتمكن مستخدم الأداة من الحكم على مدى تمثيل عينة الأسئلة للمجال الذي تقيسه.

ii. الصدق المرتبط بالمحك:

ويطلق عليه عدة تسميات منها الصدق العملي أو الواقعي؛ الصدق التجريبي، ويقصد به مجموعة الإجراءات التي نتمكن من خلالها من حساب الارتباط بين الاختبار الحالي ومحك خارجي مستقل(بشير معمرية، 2007، ص140) والمحكّات أنواع وهي(الفروق في العمر، التحصيل الدراسي، الأداء في المهنة، الأداء في التدريب، الارتباط باختبار آخر، التقديرات).

والهدف من استخدامه هو تقدير إمكانية إحلال الاختبار الجديد بإجراءاته البسيطة وغير المكلفة محل المحك؛ الذي قد يصعب الحصول عليه أو يتطلب وقتا كما يستخدم في حال توفر محك خارجي مستقل ثابت وصادق(سعيد حسن آل عبد الفتاح الغامدي، 2003، ص 30) وعليه فإن قيم معاملات الصدق المرتبطة بمحك تتأثر بعدد من العوامل يجب أخذها بعين الاعتبار عند تفسير الدرجات مثل ما ذكره(سعيد حسن آل عبد الفتاح الغامدي، 2003، ص 30-31):

1. **تجانس العينة:** يؤدي تجانس العينة إلى انخفاض معامل الصدق، ذلك أن أحد المفاهيم الهامة لصدق المقياس هو قدرته على تمييز وإظهار الفروق الفردية لدى الأفراد في مجال سمة أو قدرة معينة، هذا من جهة، ومن جهة ثانية فإن معامل الصدق هو في جوهره معامل ارتباط يتأثر كثيرا بمدى زيادة أو نقصان الفروق الفردية

2. **درجة ثبات المحك:** حيث يجب أن يكون معامل ثبات المحك عاليا لأن القيمة الحقيقية لمعامل الصدق لا يمكن أن تتجاوز مؤشرات الثبات.
3. **ذاتية المصحح:** قد تؤثر تقديرات المصححين للأفراد وبالتالي يتأثر معامل الصدق
4. **طول المقياس:** له دور كبير في زيادة معامل الصدق فنظرا لأن القيمة القصوى للصدق المرتبط بالمحك تعتمد على مؤشر الثبات فإن هذه القيمة تزداد بزيادة قيمة معامل الثبات، وزيادة قيمة معامل الثبات تعتمد على طول الاختبار أي عدد فقراته.
5. **المدة الزمنية الفاصلة بين تطبيق المقياس التنبؤي ومقياس المحك:** لها تأثيرها نظرا لأن معامل الصدق ينخفض بازدياد المدة الزمنية الفاصلة وذلك لتأثير الأخطاء العشوائية في درجات كل من المقياسين.
6. **تسرب العينة:** كما ينبغي أن نبدأ بعينة كبيرة من الأفراد تحسبا لفقدان بعضهم عند تطبيق مقياس المحك مما يؤدي إلى نقصان معامل الصدق.

وهناك طريقتان للتحقق من صدق المحك هما:

1. **الصدق التلازمي**(Congcent Validity): ويهدف إلى كشف العلاقة بين الاختبار ومحك تُجمع بياناته في وقتٍ أو قبل إجراء الاختبار الحالي.
2. **الصدق التنبؤي**(Predictive Validity): ويهدف إلى إيجاد العلاقة بين نتائج الاختبار الحالي ونتائج محك بعد أن نحصل عليها مستقبلا.

iii. صدق التكوين الفرضي:

ويقصد به مدى استيعاب المفهوم المقاس للوقائع أو الأنماط السلوكية الواقعية ويعرفه تقرير اللجنة القومية الأمريكية لمعايير الاختبارات على أنه: " يقدر صدق التكوين الفرضي بفحص أي الخصائص يقيس الاختبار بمعنى أن نحدد المفاهيم التفسيرية والتكوينات النظرية المعينة المسؤولة على الأداء على الاختبار"(بشير معمرية،2007، ص149) وهو لذلك يشمل(الفروق بين المجموعات، التغير في الأداء، المقارنة الطرفية، حساب الارتباط بين الخاصية وخاصية أخرى) وهو يقوم على التحليل النظري للمفهوم من خلال اختبار الفروض التي تقوم عليها النظرية، كما يستعان بهذا النوع من الصدق في حال كون الخاصية المراد قياسها لا يمكن ملاحظتها مباشرة، ولكنها مفهوم نظري يمكن إدراكه والاستدلال عليه من خلال بعض مظاهر السلوك وهو غالبا ما يكون مجردا.

وقد بين ساكس(Sax) وزيلر(Zeller) كما أورد في (عفاف بنت راضي مشخص اللحياني، 1429– 1430، ص87– 88) أنه يمكن إيجاز عملية التوصل إلى أدلّة تتعلق بصدق البناء في الخطوات التالية على الأقل:

1. تبرير أهمية التكوين الفرضي من الناحية التربوية والنفسية في أن يكون التكوين الفرضي معرفا تعريفا إجرائيا ويدل على سمة قابلة للقياس.
2. الاستناد إلى نظرية تربوية أو نفسية أو اقتراح نموذج منطقي يوضح المفاهيم والعلاقات القائمة بينها.
3. التمييز بين التكوين الفرضي والتكوينات الفرضية الأخرى المماثلة له.
4. التوصل إلى أدلة من مصادر متعددة باستخدام الطرق الارتباطية والتجريبية لتأكيد التكوين الفرضي، وعندما تستخدم أدوات متعددة يمكننا الحصول على الصدق التقاربي للتكوين الفرضي.
5. التوصل إلى أدلة تؤكد أن التكوين الفرضي يرتبط بعوامل وقتية أو دخيلة لكي نحصل على الصدق التمايزي للتكوين الفرضي.
6. إجراء تعديلات مستمرة في التكوين الفرضي بما يتفق والأدلة والمعلومات الجديدة المتجمعة.

يعتمد صدق الاختبار بالأساس على صدق مفرداته وذلك من خلال حساب الارتباط القائم بين الأسئلة والمحك(داخلي، خارجي) وعليه مهما كان نوع الصدق المعتمد فإنه يمكن استخراج مؤشراته حسب(عبد المنعم أحمد الدردير2006، فؤاد البهي السيد1979) والتي تتمثل أساسا في طريقة: الارتباط الثنائي الأصيل، صدق الفروق الطرفية، صدق المقارنة الطرفية

أ. طريقة الارتباط الثنائي الأصيل:

تعتمد هذه الطريقة على حساب معامل الارتباط الثنائي الأصيل لمفردات المحك(الداخلي، الخارجي) وتصلح في حال الاختبارات التي تعطي تدرجا للقدرة المقاسة أو من نوع الصواب والخطأ(الدرجات الثنائية)، وذلك من خلال حساب الارتباط القائم بين كل سؤال والدرجة الكلية للاختبار بتطبيق المعادلة التالية:

$$ر_{ث} = \frac{م_{أ} - م_{ب}}{ع} \times \sqrt{أ \times ب}$$

(فؤاد البهي السيد، 1979، ص457)

حيث:

- ر ث: معامل الارتباط ثنائي الأصيل
- م أ: متوسط الصواب
- م ب: متوسط الخطأ
- أ: نسبة الصواب
- ب: نسبة الخطأ
- ع: الانحراف المعياري لدرجات الميزان(المحك)

تعتبر هذه الطريقة من أدق الطرق المعروفة لحساب صدق مفردات الاختبار التحصيلي، لكنها تستغرق الوقت والجهد خاصة في حال طول الاختبار وكثرة عدد التلاميذ، وعليه أوجدت طرق أخرى أسرع.

ب. طريقة المقارنة الطرفية:

تقوم فكرة هذه الطريقة على تقسيم درجات المحك(الميزان) إلى مستويين ممتاز وضعيف (بناء على ترتيب الدرجة الكلية على الاختبار لكل تلميذ) لتقارن بعدها درجات السؤال بين هاتين المجموعتين(يؤخذ المستوى الضعيف والممتاز بناء على الوسيط أو نسبة 27% عادة) وتتلخص العمليات الحسابية لتقدير صدق السؤال في مقارنة معامل سهولة السؤال في الجزء العلوي بمعامل سهولته في الجزء السفلي وقد أوجد فلانجان طريقة سريعة لحساب معامل الارتباط وأوجد جداول خاصة بذلك حيث يدل السطر الأفقي الأول في جميع تلك الجداول على عدد الناجحين في السؤال من الجزء العلوي للاختبار المساوي لـ 27 % من العدد الكلي للأفراد، ويدل العمود الرأسي الأول في جميع تلك الجداول على نسبة الناجحين في الجزء السفلي للاختبار المساوي لـ 27 % من العدد الكلي للتلاميذ، وتدل الخلايا الداخلية لتلك الجداول على معاملات الارتباط.

جـــ. صدق الفروق الطرفية:

تعتمد طريقة الفروق الطرفية على نفس فكرة المقارنة الطرفية وقد وضع جونسون معادلة مختلفة عن طريقة فلانجان ولكنها تحقق نفس النتائج وتتمثل فيما يلي:

$$\text{معامل صدق السؤال} = \frac{\text{ص ع} - \text{ص س}}{0.27 \times \text{ن}}$$

(فؤاد البهي السيد، 1979، ص461)

حيث:

- ص ع: إجابات المفردة الصحيحة في القسم العلوي
- ص س: إجابات المفردة الصحيحة في القسم السفلي
- ن: عدد التلاميذ الذين أجابوا على السؤال

سادسا- صدق الاختبار كمشكل يعانيه مصمموا الاختبارات التحصيلية:

ومن الأمور التي تجعل الامتحانات أمراً هاماً نظرة الأسرة والمجتمع لهذه الامتحانات، فالمجتمع والأسرة يراقبان ما يجري داخل المدرسة بعين الفاحص والناقد، إجراءات الامتحانات ونتائجها، وبسبب ما يؤثر ذلك على مستقبل التلاميذ وعلى نظام حياتهم وبسبب ما تنفقه الأسرة أو المؤسسة التعليمية من إمكانيات ومال وجهد، ونحن قلما نجد من ينظر إلى الامتحانات نظرة محببة عادية، فهي في رأي العديد من الناس عثرة في وجه التربية والتعليم،... ثم إن الامتحانات قلما تكون وسيلة موضوعية لقياس التحصيل الدراسي، فالمعلم يضع اختباراً، ويخصص له وقتاً محدداً، ثم يستخرج النتائج ليعممها على تحصيل التلاميذ في الموضوع الذي يتناوله الامتحان ويثير هذا الموقف التقويمي ثلاث قضايا هي حسب الزراد (كما أورد في: فؤاد علي العاجز،2008)

1- أن ما يقاس لا يخضع للقياس مباشرة.
2- لا يسمح بقياس موضوعي لقدرات المفحوص.
3- أن القياس ينطوي على تعميم تكون أسسه موضوع شك وجدل وينطوي عليه آثار على سلوك التلميذ ونفسيته، والواقع أن ما يجري داخل التلميذ من تغيرات لا يمكن قياسها تماماً، وإنما نقيس من خلال الامتحان المظاهر الخارجية المتمثلة في الأداء.

وفي دراسة متعمقة حول الامتحانات في الوطن العربي وجد السنبل (كما أورد في: فؤاد علي العاجز،2008) أن:

1- الامتحان التقليدي في أية مادة لا يقيس شيئاً واحدا بل أشياء كثيرة مختلطاً بعضها ببعض، ومن القواعد المقررة عند علماء التقويم التربوي أن المقياس الذي يستعمل لقياس شيئين في آن واحد لا يقيس شيئاً.
2- ورقة الامتحان في بعض المناطق قد يصادف فيها التلميذ الحظ، وتأتي الأسئلة مما يتقنه، أو تعاكسه في امتحان آخر فتأتي الأسئلة مما لا يعرفه.
3- نظام الامتحانات الحالي يحكم على مستوى أداء التلميذ من حيث النجاح أو الرسوب ، ولكنها لا تشتمل على النواحي التشخيصية التي تبين بوضوح نواحي ضعف التلاميذ حتى يعالجها في محاولته التالية.

من هنا نستطيع القول بأن هناك العديد من العوامل المتعلقة بنظام الامتحانات السائد في بلادنا، وكلها ذات أثر عليها بحيث ينطوي إقرار صدق أي قياس على إستراتيجية شاملة يمكن تطبيقها على المقاييس المقننة والمقاييس التي يطورها المعلمون، وحسب رجاء محمود أبو علام(1978) فإنه إذا كان اهتمام التقويم هو التحصيل الشامل في مادة

معينه، فإن تطوير الاختبارات يبدأ بوضع مخطط تفصيلي يعرف بجدول المواصفات، وينبغي مقارنة هذا المخطط التفصيلي الذي يشخص المفاهيم الأساسية، ومستوى المهارات المراد تقويمها بخارطة لمجال المادة الدراسية لضمان شمول الاختبار للمفاهيم المهمة والثانوية؛ وهناك تفسيرات محتملة للأداء الخاطئ، أو الضعيف، وهي:

1. أن الاختبار صعب جدًّا بالنسبة للتلاميذ.
2. غموض التعليمات أو تشويشها.
3. قياس الاختبار لتركيبات أو مهارات أخرى.

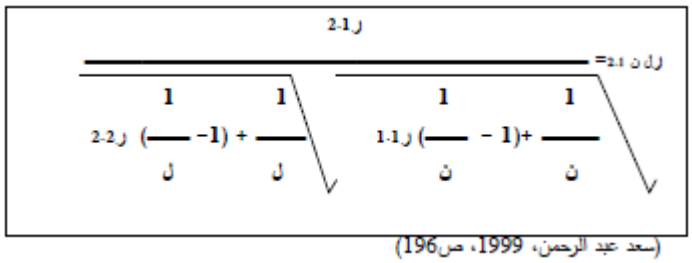
وعليه فإن إهمال تنفيذ الإجراءات التي تناولت الصدق يبعد عن تحقق سلامة الاستنتاجات المستنبطة من التقويم، ومن ثم يتم الحصول على نتائج غير صحيحة، ويرى بعض علماء القياس أن الأخطاء المتأتية من المقياس تعد من أكثر الأخطاء تأثيراً في الدرجة الملاحظة، إذ يمكن إلى حد ما السيطرة على الأخطاء المتأتية من العينة من خلال الاختيار الصحيح والتمثيل الدقيق لها للمجتمع، ويمكن تقليل الأخطاء المتأتية من الظروف التجريبية أو الظروف المحيطة بعملية القياس من خلال تثبيت هذه الظروف أو عزل تأثيرها تجريبياً أو إحصائياً، في حين مهما حاولنا أو بالغنا بدقة إجراءات بناء المقياس لا بد من وجود أخطاء في درجاته، لأنه يحتوى على مجموعة من الفقرات ويقيس عينة من السلوك الذي نستدل منه على السمة التي نريد قياسها وعليه من المتعذر أن تقيس كلّ السلوك وبشكل مباشر(حسان العمري، 2011 ، ص112–113)؛ وقد يكون ذلك نتيجة تدخل بعض العوامل ذكرها كل من (سامي محمد ملحم 2002، سعد عبد الرحمن 1999، فؤاد البهي السيد 1979) وهي على النحو التالي:

1. عوامل تتعلق بالاختبار نفسه:

- **لغة الاختبار:** إذا كانت لغة الاختبار فوق مستوى التلميذ فإن بعضهم سوف يعجز عن فهم السؤال وبالتالي لا يستطيع الإجابة عنه مما يقلل من مستواه في الاختبار
- **غموض الأسئلة:** مما يجعل التلميذ يفسرها تفسيرات متباينة ويجيب عنها إجابات خاطئة تقلل من مستواه الفعلي في الاختبار
- **سهولة الأسئلة أو صعوبتها:** تجعل التلميذ يحصل على درجات لا يستحقها فيتأثر حكمنا عليه بأنه تلميذ ممتاز أو وسط أو ضعيف، وهي لا تمثل في الحقيقة قدرة التلميذ الفعلية في الاختبار.
- **الأسئلة الموحية بالإجابة:** مما يمكّن التلميذ من الحصول على درجة أعلى مما يستحق.
- **علاقة الأسئلة بما تعلمه التلميذ فعليا:** فإذا كانت الأسئلة الاختبار هي غير ما تعلمه التلميذ تؤدي إلى حصوله على درجات لا تمثل تحصيله الفعلي وتقلل بالتالي من مستوى صدق الاختبار.
- **طول الاختبار:** يزداد صدق الاختبار تبعا لزيادة عدد أسئلته لأن ذلك الطول يضعف أثر الشوائب أو أخطاء القياس لكبر حجم عينة الأسئلة، وبذلك يزداد معامل ارتباط الاختبار بالميزان وترتفع القيمة العددية لمعامل صدق الاختبار، وهناك عدة حالات توضح علاقة صدق الاختبار بطوله.

أ- **عندما يزيد طول الاختبار "ن" مرة ويزيد طول المحك "ل" مرة:**

يمكن التعبير عن هذه العلاقة بالمعادلة التالية:



(سعد عبد الرحمن، 1999، ص196)

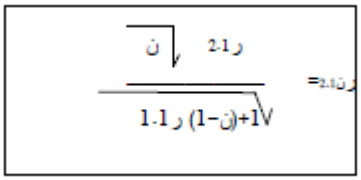
حيث:

- ر ل ن 2.1: معامل صدق الاختبار بعد زيادته ن مرة وزيادة المحك ل مرة.
- ر 2.1: معامل صدق الاختبار قبل الزيادة(معامل الارتباط بين الاختبار والمحك الخارجي).
- ر 1.1: معامل ثبات الاختبار.
- ر2.2: معامل ثبات المحك الخارجي.
- ن، ل : عدد مرات الزيادة

مثاله:

لو افترضنا أن لدينا ر 2.1(0,60) ، ر 1.1(0,70)، ر2.2(0,80)، ن(3 مرات)، ل(مرتين) من خلال التطبيق العددي في المعادلة السابقة نحصل على ر ل ن 2.1 (0,72) وعليه نلاحظ أن معامل الصدق ارتفع من (0,60) قبل الزيادة إلى (0,72) بعد زيادة الطول.

ب- **عندما يزيد طول الاختبار ن مرة ويبقى المحك الخارجي كما هو: يمكن التعبير عن هذه العلاقة بالمعادلة التالية:**



(سعد عبد الرحمن، 1999، ص196)

حيث:

- ر ن 2.1: معامل صدق الاختبار بعد زيادة طوله ن مرة
- ر 2.1: معامل صدق الاختبار قبل الزيادة.
- ر 1.1: معامل ثبات الاختبار.
- ن: عدد مرات الزيادة

مثاله:

لو افترضنا أن لدينا ر 2.1(0,60) ، ر 1.1(0,70)، ن (4 مرات) وبالتطبيق في المعادلة السابقة نجد أن ر ن 2.1(0,68)، نلاحظ أن معامل الصدق كان (0,60) فزاد إلى (0,68)

ج- **عندما يزيد طول الاختبار إلى ما لا نهاية أي يصبح ثابتا تمامـــا(ر 1.1= 1):** وفي هذه الحالة يصبح الصدق بعد الزيادة هو النسبة بين معامل الصدق القديم ومعامل الصدق الذاتي(الجذر التربيعي لمعامل الثبات) أي:

$$ر_{2.1 \infty} = \frac{ر_{2.1}}{\sqrt{ر_{1.1}}}$$

(سعد عبد الرحمن، 1999، ص197)

حيث:

- ر 2.1 ∞: معامل صدق الاختبار بعد الزيادة
- ر 2.1: معامل صدق الاختبار قبل الزيادة.
- ر 1.1: معامل ثبات الاختبار.

مثاله:

لو افترضنا أن لدينا ر 2.1 (0,60) ، ر 1.1 (0,70) بعد التطبيق العددي فـــي المعادلـــة نجـــد أن ر 2.1 ∞ (0,72) : نلاحظ أن معامل الصدق كان (0,60) فزاد إلى (0,72)

د- عندما يزيد طول الاختبار إلى ما لا نهاية ويزيد طول المحك إلى ما لا نهاية: يعبر عنها بالمعادلة التالية:

$$ر_{\infty . \infty} = \frac{ر_{2.1}}{\sqrt{ر_{1.1} \times ر_{2.2}}}$$

(سعد عبد الرحمن، 1999، ص198)

حيث:

- ر ∞. ∞: معامل صدق الاختبار بعد الزيادة
- ر 2.1: معامل صدق الاختبار قبل الزيادة.
- ر 1.1: معامل ثبات الاختبار.
- ر 2.2: معامل ثبات المحك.

مثاله:

لو افترضنا أن لدينا ر 2.1(0,60) ، ر 1.1(0,70)، ر 2.2(0,80) بعد التطبيق العددي في المعادلة نجد أن ر ∞. ∞ (0,81): نلاحظ أن معامل الصدق كان (0,60) فزاد إلى(0,81)

- **ثبات الاختبار:** يتأثر الصدق بالقيمة العددية لمعامل الثبات تأثرا مباشرا ومطردا فيزداد الصدق تبعا لزيـــادة الثبـــات الذي يتأثر هو الآخر بطول الاختبار

2. عوامل تتعلق بتطبيق الاختبار وتصحيحه:

- **العوامل البيئية:** تؤثر هذه العوامل على أداء التلميذ تأثيرا يقلل أو يزيد من مستوى صدق الاختبار مثل الحرارة، البرودة، الضوضاء، الزحام،..
- **طباعة أسئلة الاختبار:** عدم وضوح الأسئلة إضافة لوجود أخطاء مطبعية أو سوء ترتيب أسئلة الاختبار تقلل من مستوى صدق الاختبار
- **التعليمات غير الواضحة:** بحيث لا تبين للتلميذ كيفية الإجابة عن أسئلة الاختبار فيخطئ التلميذ في فهم تعليمات الاختبار ويحصل على درجات أقل مما يستحق بالفعل.
- **استعمال الاختبار في غير ما وضع له.**
- **عدم استعمال الاختبار مع الفئة التي وضع لها:** يقلل من قدرة المفحوصين على الحصول على درجات أعلى.

3. عوامل تتعلق بالتلميذ:

- **اضطراب التلميذ في الاختبار:** مما قد يشل قدرته على الإجابة ويحصل بالتالي على نتيجة لا تمثل قدراته الحقيقية.
- **التخمين أو الغش:** أو محاولات التلميذ التأثير على الفاحص بالأسلوب الذي يؤدي به إلى اختباره .
- **التباين:** من أحد المفاهيم الهامة لصدق الاختبار هو مقدرته على أن يميز بين طرفي القدرة التي يقيسها بمعنى إظهار الفروق الفردية في مجال هذه القدرة، وعليه فإن معامل صدق الاختبار يتناسب طرديا مع تباين درجات المجموعة؛ فالتباين الضعيف يقلل من أثر الصدق بينما يزيد التباين القوي من القيمة العددية لذلك الارتباط ويصل الصدق إلى نهايته الصغرى عندما يصل تباين الاختبار والميزان إلى النهاية الصغرى أيضا، أي عندما تزول الفروق القائمة بين الأفراد في درجات الاختبار ودرجات الميزان؛ مع ملاحظة أن زيادة التباين هي زيادة التباين الحقيقي الذي يؤدي إلى إظهار الفروق الفردية ويتناسب طرديا مع القيمة العددية لمعامل الصدق.

سابعا– عيوب النظرية الكلاسيكية في القياس:

تسلم النظرية الكلاسيكية في القياس بأن أداء الفرد في المقياس يعكس درجة الفرد الحقيقية في الصفة التي يقيسها والخطأ العشوائي الذي ينتشر حول هذه الدرجة، كما اتسمت النظرية التقليدية للصدق بالتركيز على التمايز الوظيفي لأنواع الصدق، أي أن لكل نوع وظيفة خاصة به تختلف عن وظائف الأنواع الأخرى من الصدق وترتب عن ذلك أن كل نوع يناسب بعض أنواع الاختبارات والمقاييس ولا يناسب الأنواع الأخرى فصدق المحتوى يناسب اختبارات التحصيل، والصدق المحكي يناسب اختبارات القدرات والاستعدادات، وصدق المفهوم يناسب اختبارات الشخصية؛ ولقد كان معظم القياس في التربية وعلم النفس قائما على النظرية الكلاسيكية إلا أن ذلك معرض لجملة من الانتقادات قللت من موضوعيته ذكرها (رجاء محمود أبو علام 2004، والسيد محمد أبو هاشم حسن) أهمها:

1. تأثر خصائص مفردات الاختبار بقدرة الأفراد المفحوصين: طبيعة العينة فإذا لم تكن مناسبة للاختبار فإن نتائجه ستكون مضللة.
2. تأثر الدرجة الكلية للفرد في اختبار ما بمفرداته: فقد يكون الاختبار سهلا جدا أو صعبا جدا بالنسبة لبعض الأفراد وتحت مثل هذه الظروف فقد يكون المقياس تقديرا ضعيفا للدرجة الحقيقية لقدرة الفرد التي نقيسها.
3. عدم توافر وحدة قياس ثابتة.
4. عدم إمكانية المقارنة بين تحصيل الأفراد في حالة استخدام اختبارات تحصيلية مختلفة.
5. تأثر ثبات الاختبار بالموقف الاختباري: طبقا للنظرية الكلاسيكية في القياس تتحدد كمية خطأ القياس في مفردات الاختبار (في مقابل خطأ القياس الذي يمكن أن ننسبه لظروف الاختبار وتقدير درجاته، وللأفراد الذين أخذوا الاختبار) بإجراء معاملات ارتباط بين أداء الأفراد في صورة مكافئة من الاختبار، ومن الناحية العملية فإنه من

الصعب جدا بناء صورة مكافئة تماما وعليه فإن معامل الثبات الذي نحصل عليه بهذه الطريقة قد لا يكون فعلا ممثلا لمعامل ثبات الاختبار.

6. صعوبة إثراء الاختبارات المعدة باستخدام النظرية الكلاسيكية بمفردات جديدة.
7. تساوي تباين أخطاء القياس لجميع أفراد العينة موضع الاختبار: تسلم النظرية الكلاسيكية في القياس أن كمية خطأ القياس واحدة من فرد لآخر من الأفراد الذين طبق عليهم المقياس، ولكن الواقع العملي يبرز أن اختبارا ما قد يكون معامل ثباته أعلى(أي خطأ قياس أقل) للأفراد في أحد مستويات القدرة التي يقيسها الاختبار من أفراد آخرين في مستوى أخر من مستويات القدرة ذاتها.

يذكر رجاء محمود أبو علام(1978) أن علم النفس التطويري نظر إلى النزعات والقدرات البشرية على أنها أصبحت متركزة في الطفولة المبكرة، وعُدّت السمات غير متغيرة بصورة عامة؛ حتى يتقدم العمر، وهذا رأي " كلارك " وفي الخمسينات تطورت معايير مهنية للاختبارات والتقنيات التشخيصية وتمت مراجعتها عام1966، وصنّفت أنواع الصدق ضمن ثلاثة أنواع وهي(صدق المحك، صدق المحتوى، صدق المقياس) أما في المنظور الحالي وعبر مرّ السنين أسهم تغيّر الأولويات الاجتماعية للتعليم، والبحث في السّمات الإنسانية، والتدخلات التربوية في توسيع النظرة إلى الصدق، ففي الدرجة الأولى لم يعد يُنظر إلى التعليم الرسمي كنظام جامد يمكن من خلاله لعدد قليل من التلاميذ فقط الحصول على مستويات عالية، وقد أشارت الدراسات البحثية إلى أهمية تأثير البيئة في التحصيل، فضلًا عن أن الاستخدام المتزايد للاختبارات فيما يتعلق بالقرارات المصيرية، مثل: الترقية إلى الفصل التالي، والحصول على شهادة ثانوية، قد أدت أيضا إلى طُعونات قضائية في بعض هذه الاختبارات، وقد أسهمت هذه الوقائع في تركيز الاهتمام حاليًا على معنى درجات الاختبار، وتجسد معايير 1985 للاختبارات التغير في الاهتمام من خلال تعريف الصدق بوصفه ملائمة، وجديدة، وفائدة.

ثامنا- نظرية الصدق الحديثة:

غيرت نظرية الصدق الحديثة بعض التصورات تغييرا جذريا، إن للمفهوم الحالي للصدق سمات رئيسية تدعي:

1. أن الفكرة التي تولدت منذ ثلاثين عاما بأن هنالك ثلاثة أنواع منفصلة من الصدق لا يمكن أن تلبي المتطلبات الحالية للاختبارات، وبدلًا من ذلك فإن صدق المحتوى والمحك ما هما إلا خيطان من جديلة الصدق.
2. أن الصدق عملية تكاملية، وليس أجزاء، فضلًا عن أن الإجراءات المتعددة المتخذة تنفذ من مراحل مختلفة في عملية بناء التقويم.
3. إن الصدق يعد مفهومًا ديناميا، وليس مفهومًا ثابتًا.

وتتلخص أهم تطورات أو جوانب الجدة في مفهوم الصدق فيما يلي حسب (أمحمد بوزيان تيغزة):

1. أصبحت نظرية الصدق الحديثة تركز على مدى ملاءمة عملية تأويل درجات المقياس أي أن الصدق يتمثل في طبيعة الأدلة التي تقدم للدلالة على مدى دقة تأويل درجات أداء المفحوصين على الاختبار أو طريقة قراءتها ونبذت تصور الصدق باعتباره صفة أو خاصية للمقياس.
2. الأخذ بالنظرة الواحدية للصدق ونبذ فكرة تعدد أنواع الصدق فأنواع الصدق كلها أُدمجت تحت مظلة واحدة سميت بصدق التكوين الفرضي أو البنائي أو صدق المفهوم، وأمسى مفهوم " صدق المفهوم" يمثل كل مجال الصدق بشتى أنواعه التقليدية، ولا يمثل فقط أحد أنواعها.
3. إن توحيد مفهوم الصدق أو مجاله لا يعن أن الأدلة الدالة عليه متماثلة أو واحدة.

4. أضحى البعد القيمي الاجتماعي لعملية القياس يشكل جزءً لا يتجزأ من مفهوم الصدق ولذلك أضيف بعد جديد للصدق تمثل في المآل أو المترتبات أو النتائج الاجتماعية لعملية القياس؛ وسواء أكانت تلك النتائج نتائج إيجابية متوقعة، أو نتائج سلبية غير متوقعة واتخذ هذا البعد الجديد للصدق أحيانا مسمى" صدق المآل أو المترتبات.
5. لم يعد ينظر للثبات كمجال قائم بذاته ومستقل عن الصدق رغم العلاقة التي تربط بينهما وإنما أصبح ينظر إليه كنوع من البينات(بينة من بينات الصدق)، أو جانب أو وجه من أوجه الصدق.

ويضيف أمحمد بوزيان تيغزة القول بأنه على الرغم من أن صدق المفهوم يعتبر تجديدا بل تطوّرا جوهريا في التنظير لنظرية الصدق الكلاسيكية في الخمسينيات إذ استهدف ظهوره التخفيف من غلواء الإجراءات الأمبيريقية الآلية التي قامت عليها ممارسات الصدق المحكي، وأعاد للتأصيل التنظيري، غير أن صدق المفهوم عُومل كنوع من أنواع الصدق من جهة، وبأنه نوع مكمل يستعمل عند عدم ملاءمة الأنواع الأخرى لتقدير الصدق، فصدق المضمون، والصدق المحكي بنوعيه(الصدق التلازمي والصدق التنبؤي) يحظى بأولوية الاستعمال ولا يلجأ إلى صدق المفهوم إلا إذا استحال تحديد المجال السلوكي للمفهوم، أي استحال استعمال صدق المحتوى، وأيضا عند استحالة العثور على محك مناسب للمفهوم؛ فميهل وكرونباخ (1956) وهما مهندسا مصطلح صدق المفهوم يريان على غرار الدليل الإرشادي الأول للقياس، بأن الباحث الذي يعتقد أن المحكّات المتوفرة غير صادقة أو غير مناسبة، أو تعذر وجوده فيلجأ بدل ذلك إلى صدق المفهوم للخروج من هذا المأزق، أما بالنسبة لصدق المحتوى فالقبول بمجتمع المحتويات الذي يحدد المتغير الذي يراد قياسه يعتبر أمرا جوهريا، وفي ظل غياب المحك المناسب، ومجال المحتويات الذي يفترض أن يحدد الصفة التي يراد قياسها، لا بد من اللجوء إلى توظيف صدق المفهوم.

غير أن هذا المنظور لصدق المفهوم باعتباره نوعا مكملا لأنواع الصدق الأخرى التي تتمتع بأوّلية الاستعمال تغير تغيّرا جذريا انطلاقا من الثمانينيات إلى الوقت الحاضر، ويلاحظ كرونباخ بحق بأن" المفهوم "أو" البناء الفرضي" يهدف إلى تقدير الصدق، أو إنتاج البينات والشواهد التي تقيم الحجة على توفر الصدق مادمنا نتصور بأن الصدق يُعنى أساسا بأوجه التأويل أو التفسير أو الاستدلالات القائمة على درجات المقياس ولا يُعن بالدرجات في حد ذاتها أو المقياس في حد ذاته، فالمجال الواسع لعناصر المحتوى بالنسبة لصدق المحتوى والوقائع وأبعاد السلوك الدالة على المحك بالنسبة للصدق المحكي أو الصدق المرتبط بالمحك لا تنشأ في فراغ، وإنما تتأسس هذه المحتويات والوقائع وأبعاد السلوك بناءً على المفهوم أو التكوين الفرضي الذي يراد قياسه، فالإطار المرجعي لعناصر مجالات المحتوى وللأبعاد المختلفة للمحكّات يتجلى في المفهوم أو التكوين الفرضي.

وينبه أمحمد بوزيان تيغزة أن مهندسي فكرة وحدة الصدق وأن هذه الوحدة يحتضنها صدق المفهوم أعمال ميسيك (Messick, 1994) التي تمحورت أساسا حول استجلاء مفهوم شامل للصدق وتجديد نظرية الصدق التقليدية، فلما كان الصدق في تصور ميسيك يُعنى بدمج متكامل لمصادر البيّنات التي تستهدف تفسير وتأويل وتبين دلالة درجات المقياس سواء أكان مصدر هذه البينات محتوى المقياس (الذي يشكل مناط صدق المحتوى) أم كان مصدرها البينات المرتبطة بالمحكّات (التي تشكل مناط الصدق المحكي) فإن صدق المحتوى وصدق المحك يشكلان بالتالي أجزاء أو أقساما تنضوي تحت مظلة صدق المفهوم؛ إن الصدق مفهوم موحد ولما كانت قوته المحركة تتمثل في مصداقية ودلالة عمليات التأويل لدرجات المقياس وما يترتب عن ذلك من نتائج استعمال المقياس، أو نتائج اتخاذ القرارات القائمة على طبيعة تأويل درجات المقياس، فإن هذا كله (أي تأويل الدرجات، ونتائج الاستعمال أو نتائج اتخاذ القرار في ضوء دلالة الدرجات) هو ذاته صدق المفهوم؛ غير أن صدق المفهوم باعتباره مفهوما موحدا للتصنيف الرباعي للصدق الذي كرسه الدليل الإرشادي الأول للقياس(صدق المحتوى، والصدق التنبؤي، والصدق التلازمي، وصدق المفهوم) وموحدا للتصنيف الثلاثي الذي كرسه الدليل الإرشادي الثاني للقياس(صدق المحتوى، والصدق المحكي، وصدق المفهوم) أثار

إشكالية طرق البرهنة على صدق مفهوم درجات المقياس باعتبار "صدق المفهوم" مفهوما موحدا، ما مصادر البيّنات التي توظف للدلالة على صدق المفهوم، ما نوع الأدلة والشواهد الممكنة التي تفيد في بناء الحجة على وجود صدق المفهوم؟

إن النظرة إلى الصدق باعتبار مفهوما موحدا أي صدق المفهوم فإن وحدته لا تتنافى مع وجود أوجه أو أبعاد أو جوانب متمايزة له تفيد في توجيه طرق تقديره، واسترشادا بأحدث دليل إرشادي للقياس الذي صدر (1998) وبتحليلات المتخصصين في نظرية الصدق (ولاسيما أعمال ميسيك Messick) يمكن تلخيص طبيعة البينات على تعددها وتباينها وتنوعها في ستة أصناف أو فئات كبرى وهي حسب (أمحمد بوزيان تيغزة):

1. البينات القائمة على محتوى المقياس
2. البينات القائمة على عمليات أو سيرورات الاستجابة
3. البينات القائمة على البنية الداخلية لأدوات القياس
4. البينات القائمة على العلاقات بمتغيرات أخرى أو بينة البنية الخارجية.
5. البينات القائمة على الثبات.
6. البينات القائمة على نتائج القياس ومترتباته أو تبعاته.

تاسعا- بدائل النظرية الكلاسيكية في القياس:

جرت محاولات عديدة من علماء القياس النفسي والباحثين فيه لإيجاد وسائل وتقنيات جديدة واعتماد إجراءات دقيقة في بناء المقاييس النفسية، غايتها جميعاً الحد أو تقليل الأخطاء المتأتية من المقياس، ليصلوا إلى مستوى من الدقة التي وصل إليها الباحثون في العلوم الفيزيائية والطبيعية، إلاَّ أن ذلك قد يبدو صعباً لوجود مشاكل وصعوبات في قياس الظواهر أو الخصائص النفسية وتكميمها، ومن هذه الصعوبات أن الظواهر النفسية مفاهيم مجردة، أي ليس لها وجود مادي(حسان العمري، 2011، ص 112-113) وعليه فقد ظهرت عدة وسائل جديدة لتقويم التحصيل الدراسي تدّعي اتسامها بالموضوعية منها نظرية السّمات الكامنة، وتقوم هذه النظرية على افتراض أساسي مفاده أن هناك سمات أو خصائص معينة يشترك فيها جميع الأفراد ولكنهم يختلفون في مقدارها، وبالرغم من كونها غير منظورة ولا يمكن قياسها بطريقة مباشرة إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحظ للفرد المتمثل في استجابته على مفردات الاختبار وهذا ما يبرر تسميتها بنظرية السمات الكامنة، وعليه فإن استخدام هذه النظرية يقوم على المسلمات التالية حسب رجاء محمود أبو علام(2004):

1. يمثل أداء الفرد في الاختبار قدرة وحيدة.
2. يختلف أداء الأفراد في الاختبار طبقا لاختلاف كمية القدرة التي يمتلكونها.
3. يمكن تمثيل العلاقة بين متغيرات القدرة والأداء في المفردة بدالة رياضية.

لقد جاءت هذه النظرية بمنطلقات تتصدى للكثير من السلبيات التي تشوب النظرية الكلاسيكية في القياس ويؤكد الشرقاوي وآخرون أن أهمية النظرية المعاصرة في القياس تأتي من تركيزها على المستوى الذي يصل إليه أداء الفرد وتقدير اكتسابه أو تحصيله، وهذا هو الاهتمام المتعارف عليه للمقاييس والاختبارات التربوية المعاصرة؛ ومن التطبيقات التربوية والعملية لنظرية السمات الكامنة بناء بنوك للأسئلة Items Bankius وبناء الاختبارات محكية المرجع، وتكافؤ درجات الاختبارات(Test score Equating) التي تسمح بتوحيد مجاميع عديدة من الفقرات باختبار واحد، والاختبارات المفصلة، Tailored Testing؛ وبخصائص قياسية موضوعية ، فضلا عن إمكانيتها الربط بين خصائص الفقرة وخصائص الفرد نفسه مما يساعد ذلك على تشخيص مواطن القوة والضعف لمستوى التلميذ الدراسي مما يتيح

للمختصين تقديم النصح والإرشاد والتوجيه للمتعلمين (محمد أنور محمود السامرائي، أحمد محمد شاكر الخفاجي، 2012، ص 968- 969)

- **خلاصة:**

تبعث غالبية الأسر الجزائرية بأبنائها فلذات أكبادها إلى المدارس طمعا في احتلالهم لأعلى المراتب وتخطيهم مراحل التعليم بجدارة، لكن الواقع يكشف عن وجود ليس محطات فقط بل وعوائق من نوع ما قد تقف حائلا دون الترقي من مستو لآخر، ولعل من هذه العوامل هو نظام الامتحانات المدرسية وما يشوبها من ظروف بداية من إعدادها إلى تطبيقها وأخيرا تصحيحها ومن ثم اتخاذ قرارات تربوية بناء على نتائجها هاهنا يمكن أن نصطدم بنتائج غير متوقعة فنعود باللوم أولا على التلميذ، لكن الدراسات التي تناولت الامتحانات تشير اليوم وبحدة إلى أن هناك عوامل خاصة بالامتحانات في حد ذاتها هي التي تؤدي إلى حرمان عدد غير قليل من التلاميذ بالفوز بفرحة النجاح؛ ولعله وعندما نذكر أهم شروط الاختبار الجيد سيتصدرها مفهوم صدق الاختبار الذي يتعلق بنتائج الاختبار وليس بالاختبار فقط، وعليه فقد تطور المنظور التعريفي الكلاسيكي من التركيز المطلق على المعيار بحيث أن الارتباط بين المقياس والمعيار هو الذي يكوّن الصدق، والتركيز أيضا على المحتوى بحيث ينظر إلى الصدق باعتباره يتكون من عناصر محتوى الاختبار التي يجب أن تكون ممثلة لمجال أوسع من المحتويات ثم التركيز على وظيفة المقياس بحيث أن الوظيفة أو الهدف، أو الغرض الذي وضع من أجله المقياس هو الذي يصنع الصدق ثم إلى التركيز على الدلالة النظرية التي تفسر أو توضح المفهوم الذي يقيسه الاختبار.

ولقد ظهرت بوادر التغير في التصور والتعريف لمفهوم الصدق في التعابير اللغوية التي صيغت بها تعريفات الصدق انطلاقا من الدليل الإرشادي الأول للقياس وانتهاء بالدليل الإرشادي الخامس والأخير للقياس الذي ظهر سنة 1999 ولكن ما زال النموذج الثلاثي للصدق سائدا سيادة تامة في التدريس والبحث والتأليف لحد الآن، ولقد أشار كرونباخ (1971) بأن عملية تقدير الصدق لا تتعلق إطلاقا بإيجاد صدق الاختبار أو صدق أدوات القياس، بل الشيء الذي يراد إيجاد صدقه هي التأويلات أو معاني الدرجات وكذلك اتخاذ القرارات المتعلقة بطريقة استعمال الأدوات المنبثقة عن عملية تأويل الدرجات.

- **قائمة المراجع:**

1. أمحمد بوزيان تيغزة(2008). نظرية الصدق الحديثة ومتضمناتها التطويرية لواقع القياس. ندوة علم النفس" علم النفس والتنمية الفردية والمجتمعية".جامعة الملك سعود.كلية التربية. قسم علم النفس.
ksu.edu.sa/sites/.../أ.د.%20أمحمد20%تيغزة.pdf
2. السيد محمد أبو هشام حسن. التوجهات المستقبلية للتقويم النفسي والتربوي وتطبيقاتها في مجال التربية الخاصة.أطفال الخليج العربي:www.gulfkids.com/ar/index.php?action=show
3. بشير معمرية(2007).القياس النفسي وتصميم أدواته للتلاميذ والباحثين في علم النفس والتربية.(ب- ط)، منشورات الحبر.الجزائر.
4. رجاء محمود أبو علام. (1978). قياس وتقويم التحصيل الدراسي.دار القلم. الكويت
5. (2004). مناهج البحث في العلوم النفسية والتربوية.ط4.دار النشر للجامعات.القاهرة
6. حسان العمري. 2011. أثر ترتيب الفقرات في المقاييس النفسية على التجانس الداخلي للمقاييس وخصائص فقراته. مجلة جامعة دمشق:27 (3+4): 111- 129
www.damascusuniversity.edu.sy/mag/.../111-129.pdf

7. خالد بن حسن شيبان التميمي(1999). " أثر كل من نوع المحكم وطول الاختبار على تحديد درجة القطع لاختبار محكي المرجع يقيس الكفايات الرياضية في العمليات الحسابية على الأعداد بالصف السادس الابتدائي بمدينة جدة". رسالة ماجستير(منشورة). كلية التربية. جامعة أم القرى. المملكة العربية السعودية
أثر20%كل20%من20%نوع20%المحكم20%وطول20%الاخ /.../www.abegs.org
8. سعد عبد الرحمن(1999). القياس النفسي(النظرية والتطبيق).ط3.دار الفكر العربي- القاهرة
9. سعيد حسن آل عبد الفتاح الغامدي.(2003). "مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تغاير بدائل الاستجابة والمرحلة الدراسية".رسالة ماجستير(منشورة). كلية التربية.جامعة أم القرى. المملكة العربية السعودية. www.gulfkids.com/ar/index.php?action...r...
10. سامي محمد ملحم(2002). القياس والتقويم التربوي في التربية وعلم النفس.ط 2. دار المسيرة للنشر والتوزيع.عمان- الأردن
11. صفوت فرج.(2000). القياس النفسي.ط4. مكتبة الأنجلو المصرية- القاهرة
12. عبد المنعم أحمد الدردير.(2006). الإحصاء البرامتري واللابرامتري في اختبار فروض البحوث النفسية والتربوية والاجتماعية. ط1. عالم الكتب. القاهرة
13. عفاف بنت راضي مشخص اللحياني(1429- 1430ه). "أثر بعض طرق تقدير الدرجات للمفردات على ثبات وصدق درجات اختبار تحصيلي في الرياضيات ذي الاختيار من متعدد لدى تلميذات الصف الأول ثانوي بمكة المكرمة" . رسالة ماجستير(منشورة). . كلية التربية. جامعة أم القرى . المملكة العربية السعودية
أثر20%بعض20%طرق20%تقد /www.jarwan-center.com/.../universitytheses
14. عماد عبد الرحيم الزغلول، شاكر عقلة المحاميد(2007). سيكولوجية التدريس الصفي.ط1.دار المسيرة للنشر والتوزيع والطباعة.عمان- الأردن
15. فؤاد علي العاجز. (يونيه 2008). دراسة تقويمية لطبيعة إجراء إمتحانات الجامعة الإسلامية من وجهة نظر طلبتها. مجلة الجامعة الإسلامية(سلسلة الدراسات الإنسانية): 16(02): 377-409
quality.iugaza.edu.ps/Files.../634700293261293244....
16. محمد أحمد محمد إبراهيم غنيم. (أبريل2003). الاتجاهات الحديثة في بحوث مشكلات تقويم التحصيل الدراسي. بحث مقدم إلى اللجنة العلمية الدائمة لعلم النفس التربوي والصحة النفسية. ..www.gulfkids.com/ar/index.php?action...r
17. محمد أنور محمود السامرائي، أحمد محمد شاكر الخفاجي. (1433ه- 2012م).بناء اختبار تحصيلي محكي المرجع في مادة علم النفس الخواص للتلاميذ أقسام العلوم التربوية والنفسية, مجلة الأستاذ: (203): 964- 1002
www.iasj.net/iasj?func=fulltext&aId=71938
18. محمد حسين سعيد حسين. تطوير أساليب التقويم ضرورة حتمية لضمان جودة المؤسسات التعليمية. ورقة مقدم إلى المؤتمر السنوي الثالث عشر للجمعية المصرية للتربية المقارنة بعنوان" الاعتماد وضمان جودة المؤسسات التعليمية" أيام 29-30 يناير 2005 www.gulfkids.com/ar/book34-1550.htm
19. فؤاد البهي السيد. (1979).علم النفس الإحصائي وقياس العقل البشري. ط3. دار الفكر العربي. القاهرة
20. نبيل عبد الهادي(2002). مدخل إلى القياس والتقويم التربوي واستخدامه في مجال التدريس الصفي. ط2، وائل للنشر والتوزيع.عمان.
21. يوسف قطامي، ماجد أبو جادو، نايفة قطامي(2000). تصميم التدريس. ط1. دار الفكر العربي للطباعة والنشر.عمان- الأردن