

معامل حساسية المفردة في الاختبارات محكية المرجع (دراسة تحليلية)

Item Sensitivity Index in Criterion-Referenced Tests (Analytical Study)

إعداد: الباحثة/ نوال بنت علي اليحياي

طالبة دكتوراه في القياس والتقويم بجامعة السلطان قابوس، سلطنة عُمان

Email: nawal.alyahyai@gmail.com

الأستاذ الدكتور/ حسين بن علي الخروصي

أستاذ القياس والتقويم بجامعة السلطان قابوس، سلطنة عُمان

المخلص:

هدفت هذه الدراسة إلى إلقاء الضوء على معامل حساسية المفردة في الاختبارات محكية المرجع كأحد أهم الأسس التي ينبغي الاعتماد عليها في اختيار المفردات في الاختبارات محكية المرجع، وقد تم استعراض الفرق بين هذا النوع من الاختبارات وبين الاختبارات معيارية المرجع للتأكيد على أن أسس اختيار المفردات في كلا النوعين ينبغي أن يكون مختلفا، وقد اعتمدت الدراسة على المنهج النظري أو التحليلي أو ما يسمى بمنهج مراجعة الأدبيات حيث تم تحليل الأدبيات العربية السابقة التي تناولت موضوع الاختبارات محكية المرجع سواء المتعلقة ببنائها أو تلك التي تناولت الطرق الإحصائية المستخدمة في هذه الاختبارات كالمقارنة بين طرق تقدير الثبات المختلفة أو المتعلقة بدرجات القطع أو بمعاملات حساسية المفردة، وقد تم ملاحظة أن هناك ندرة في الدراسات التي تركز على تناول موضوع معامل حساسية المفردة في الاختبارات محكية المرجع، كما تم ملاحظة أن بعض الدراسات المتعلقة ببناء الاختبارات محكية المرجع لا تشير إلى هذا المعامل أساسا، وبعضها يستخدم معامل التمييز بنفس الطريقة التي يستخدم بها في الاختبارات معيارية المرجع، مما يدل على وجود خلل في بناء هذا النوع من الاختبارات. وقد أوصت الدراسة الحالية بضرورة الالتفات والتركيز على الاختبارات محكية المرجع باعتبارها مكملة للاختبارات المعيارية وتعطي عمقا وفهما أوسع لمشكلات التعلم لدى الطلاب، كما أوصت بضرورة إجراء المزيد من البحوث والدراسات الإحصائية المتعلقة بطرق حساب معامل حساسية المفردة لنشر ثقافة أوسع حول هذا المفهوم في الميدان التربوي العربي، وبالتالي تطبيق أفضل للاختبارات المحكية؛ وتقديم دورات وورش تدريبية للمعلمين ولمعدي الاختبارات لتزويدهم بمهارات بناء وتحليل الاختبارات محكية المرجع؛ وتضمنين مناهج القياس والتقويم في العالم العربي لمساقات عملية في بناء وتحليل الاختبارات محكية المرجع.

الكلمات المفتاحية: معامل حساسية المفردة، الاختبارات محكية المرجع، الاختبارات معيارية المرجع.

Item Sensitivity Index in Criterion-Referenced Tests (Analytical Study)

Abstract

This study aimed to shed light on the item sensitivity index in criterion-referenced tests as one of the most critical foundations to rely on when selecting items for such tests. It reviewed the differences between this type of tests and norm-referenced tests to emphasize that the principles for item selection should differ between the two types. The study adopted the theoretical or analytical approach, also known as the literature review method, analyzing previous Arabic studies that addressed criterion-referenced tests. These included studies focused on their construction or those examining the statistical methods used in such tests, such as comparing methods for estimating reliability, cut scores, or item sensitivity indices. It was observed that there is a scarcity of studies focusing specifically on the topic of item sensitivity indices in criterion-referenced tests. Moreover, some studies on the construction of such tests do not mention this index at all, while others use the discrimination index in the same manner as it is used in norm-referenced tests, indicating a flaw in the development of this type of test. The current study recommended paying greater attention to criterion-referenced tests, considering them complementary to norm-referenced tests, as they provide deeper insights into students' learning problems. It also called for more research and statistical studies on methods for calculating the item sensitivity index to promote a broader understanding of this concept in the Arab educational field, thereby enhancing the application of criterion-referenced tests; and providing training courses and workshops for teachers and test developers to equip them with skills in constructing and analyzing criterion-referenced tests; and including practical courses in constructing and analyzing criterion-referenced tests in measurement and evaluation curricula in the Arab world.

Keywords: Item Sensitivity Index, Criterion-Referenced Tests, Norm-Referenced Tests.

1. المقدمة:

تعد الاختبارات أحد الأدوات الهامة التي يتم الاعتماد عليها في تقويم تعلم الطلبة، ومدى تحقيقهم للأهداف المرجوة من المنهج التعليمي، كما أنها وسيلة لإصدار أحكام يتم على أساسها انتقال الطلبة من مرحلة تعليمية إلى مرحلة أخرى، كما تستخدم أنواع مختلفة من الاختبارات لتقويم مدى جودة النظام التعليمي في بلد ما من خلال مقارنة نتائج الطلبة في هذا البلد مع نتائج طلبة آخرين في بلدان أخرى كما هو الحال في الاختبارات الدولية.

وعرف عودة (1985) الاختبارات بأنها طريقة منظمة لتحديد درجة امتلاك الفرد لسمة معينة من خلال إجابات الطالب عن عينة من المثبرات التي تمثل السمة. وتنقسم الاختبارات من حيث تفسير نتائجها إلى اختبارات محكية المرجع (Criterion-Referenced Tests) واختبارات معيارية المرجع (Norm-Referenced Tests). ويعتبر كلا النوعين أدوات مهمة في العملية التعليمية، واختيار النوع المناسب يعتمد على الأهداف التعليمية المرجوة.

ويعد التعلم الإيقاني أحد الوسائل الهامة للإصلاح في المجال التربوي، فهو يمثل استراتيجية قابلة للتكامل مع استراتيجيات التعلم الأخرى، ويعد التقويم التشخيصي المعتمد على الاختبارات محكية المرجع عماد التعلم الإيقاني، وعلى الرغم من ذلك، فإنه يلاحظ ندرة توافر الاختبارات التشخيصية محكية المرجع في الوطن العربي، وقد تعزى ندرة توافر الاختبارات التشخيصية محكية المرجع إلى ندرة الخبرات العربية في مجال تطوير هذا النوع من الاختبارات، من حيث معرفة كيفية بنائها، والتحليلات الإحصائية المناسبة لها (الحموري والكحلوت، 2009).

وتعد المفردات اللبنة الأساسية للاختبارات، لذا تعد مهارة صياغتها أحد أهم المعايير التي ينبغي الالتزام بها عند إعداد أي اختبار، كما أن تحليل المفردات ينبغي أن يكون على قدر كبير من الدقة وذلك للوصول إلى نتائج موثوقة بناء على درجات الاختبار المكون من هذه المفردات. ونظرا للعلاقة الوثيقة بين الاختبارات محكية المرجع وعملية التعليم، فقد اهتم خبراء القياس والتقويم بربط أساليب تحليل مفردات هذه الاختبارات بعملية التعليم من أجل تحسين نوعية هذه المفردات واختيار أكثرها حساسية وتأثرا بالتغييرات التي تطرأ على أداء المتعلمين بعد تلقيهم التعليم (علام، 2001).

ويرى عودة (1998) أن فاعلية المفردة تقاس بمدى انسجامها مع الهدف أو لا ثم بدرجة حساسيتها لعملية التعليم ثانياً. ففي الاختبارات معيارية المرجع يتم استبعاد المفردات ذات التمييز المنخفض جداً أو غير المميزة كون هذه الاختبارات تسعى لإيجاد الفروق بين الأفراد، بينما في الاختبارات محكية المرجع فإن هذا النوع من المفردات قد يشكل أهمية وقيمة لأغراض التشخيص، فإذا لم يتمكن الجميع أو أغلب الطلاب من الإجابة عليها قبل التدريب و ثم تمكنوا جميعاً أو أغلبهم من الإجابة عليها بعد التدريب فهذه المفردات في هذه الحالة لا يمكن استبعادها في الاختبارات محكية المرجع مع أنها غير مميزة باستخدام طرق معامل التمييز في الاختبارات معيارية المرجع.

ومن هذا المنطلق فإن اختيار المفردات في الاختبارات محكية المرجع يعتمد على نهج مختلف عنه في الاختبارات معيارية المرجع، فلا يمكن حساب معامل التمييز للمفردة في هذا النوع من الاختبارات ولذلك يوجد بدلا منها ما يسمى معامل حساسية المفردة والذي يساعد في معرفة المفردات التي تميز بين مجموعة الأفراد قبل التدريب وبعد التدريب مثلا، وبالتالي ما هي المفردات الأكثر أهمية في التشخيص الأولي.

ونظرا لأهمية الاختبارات محكية المرجع في تحسين عملية التقويم، وندرة وجودها في الميدان التربوي في الوطن العربي والذي قد يكون سببه قلة الخبرة فيما يتعلق بنائها وتحليلها، فقد جاء هذا البحث الذي يتناول الاختبارات محكية المرجع والذي

سيتم التركيز فيه على معامل حساسية المفردة في الاختبارات محكية المرجع وذلك لافتقار الكثير من الاختبارات في هذا الجانب من تحقيق هذا المعامل بالطرق المناسبة مما يدل على وجود مشكلة في بناء هذه الاختبارات، فمن خلال تتبع الكثير من الدراسات التي تناولت بناء الاختبارات محكية المرجع نجد أن بعضها لا يشير إلى هذا المعامل، وبعضها يستخدم نفس الطرق المتبعة في الاختبارات معيارية المرجع ونفس المعايير المستخدمة فيها للحكم على المفردة مع أن كلا النوعين (الاختبارات المعيارية والمحكية) يختلفان في فلسفة كل منهما وفي أهداف كل منهما.

1.1. أهداف الدراسة:

- 1- تهدف هذه الدراسة النظرية إلى تسليط الضوء على أحد الأدوات المهمة في القياس التربوي وهي الاختبارات مرجعية المحك من حيث التركيز على القواعد الصحيحة التي يجب الالتزام بها عند انتقاء مفردات هذه الاختبارات، وبالأخص معامل حساسية المفردة الذي من خلاله يتم التعرف على قابلية قياس المفردة لتعلم الطلاب.
- 2- كما تهدف إلى التركيز على وجود الفروق بين الاختبارات محكية المرجع ومعيارية المرجع للتأكيد على الاختلافات في طرق تحليل المفردات بين نوعي الاختبار.

2.1. أهمية الدراسة:

تأتي أهمية الدراسة النظرية من كونها من الدراسات النادرة في المكتبة العربية في حدود علم الباحث، حيث لا توجد مراجع عربية وفيرة في موضوع معامل حساسية المفردة، كما تأتي أهميتها العملية من كونها تلفت الانتباه إلى ضرورة البناء الصحيح للاختبارات محكية المرجع من خلال اختيار المفردات الملائمة والتي تشكل لبنات الاختبار الأساسية، والتي يساعد على اختيارها معامل حساسية المفردة.

3.1. منهج الدراسة:

اعتمدت الدراسة على المنهج النظري أو التحليلي أو ما يسمى بمنهج مراجعة الأدبيات حيث تم تحليل الأدبيات العربية السابقة التي تناولت موضوع الاختبارات محكية المرجع سواء المتعلقة ببنائها أو تلك التي تناولت الطرق الإحصائية المستخدمة في هذه الاختبارات كالمقارنة بين طرق تقدير الثبات المختلفة أو المتعلقة بدرجات القطع أو بمعاملات حساسية المفردة.

4.1. مصطلحات الدراسة:

معامل حساسية المفردة: معامل حساسية المفردة في الاختبارات محكية المرجع يشير إلى قدرة المفردة أو السؤال في الاختبار على تمييز مدى تحقيق المتعلمين للمستوى المطلوب من الأداء أو إتقان المهارة المستهدفة. يتم استخدام هذا المعامل لتقييم ما إذا كانت مفردات الاختبار تساعد فعلياً في قياس تحقيق الأهداف التعليمية المحددة مسبقاً ويختلف معامل حساسية المفردة في الاختبارات محكية المرجع عن معيارية المرجع، حيث تركز الأولى على مقارنة أداء المتعلم بالمحك المطلوب، وليس بأداء الآخرين. ومن ثم، يساعد هذا المعامل في تصميم اختبارات دقيقة وموجهة لتحليل وتحسين أداء المتعلمين، مع تحديد الصعوبات التي تواجههم لتصميم برامج علاجية مناسبة (طيفور وآخرون، 2003).

الاختبارات محكية المرجع: يعرف الاختبار محكي المرجع بأنه الاختبار الذي يقيس أداء الفرد بالنسبة إلى مستوى مطلق للأداء دون الرجوع إلى الآخرين وفي هذه الحالة يمكن أن نحصل على الفروق الفردية في أداء الفرد بما حققه وما لم يحققه من الاختبار أي الفرق بين الفرد ونفسه (مراد وسليمان، 2012).

5.1. حدود الدراسة

تحددت بمراجعة عينة عشوائية من الدراسات العربية التي تناولت موضوع الاختبارات محكية المرجع في العشرين سنة الماضية (2003-2024).

2. الاختبارات محكية المرجع

اعتبر علام (2000) الاختبارات محكية المرجع أكثر أنواع الاختبارات التحصيلية الملائمة لقياس تحصيل الطلبة، لأنها تتمتع بدقة فائقة تمكن المعلم من قياس الكفايات أو المهارات المحددة وملاحظتها بشكل مباشر، ويعرف عودة (2010) الاختبار محكي المرجع بأنه ذلك الاختبار الذي يقارن أداء الطالب بمستوى أداء معين يجرى تحديده مسبقاً بصرف النظر عن أداء المجموعة. ويستخدم هذا النوع من الاختبارات في الكثير من المجالات والأغراض وغالباً ما يستخدم في اتخاذ القرارات من قبل المعنيين كالمعلمين وغيرهم بشأن المعارف والمهارات والقدرات التي يمتلكها المتعلم والمتعلقة بنتائج برنامج تعليمي مثلاً أو ما شابه، وذكر علام (2001) أن الاختبار محكي المرجع يتطلب تحديد مستويات مسبقة للأداء كأن يجب الطالب على 80% من أسئلة الاختبار على الأقل، وذلك يفيد في تحديد مستويات إتقان كل فرد لأهداف معينة مرتبطة بمحتوى دراسي أو تدريبي. وذكر (Mano 2017) أن الاختبارات محكية المرجع تساعد المعلمين في تقييم الطلبة وتقديمهم، وكذلك في التخطيط للتدريس، ومن مميزاتها أن مدى التباين (الفروق الفردية) أو الاختلافات يصبح أقل ما يمكن أن يكون عليه عندما يتقن (يتمكن) الطلبة جميعاً العناصر الأساسية للموقف التعليمي، كما أن تفسير درجة الطالب يتم بالرجوع إلى درجة قطع، وليس بالرجوع إلى درجات الأقران في الصف ذاته كما هو الحال في الاختبارات معيارية المرجع (عليجات، 2022). وتعرف درجة القطع في الاختبارات محكية المرجع بأنها نقطة على المتصل يصنف عندها الحد الأقل (الأدنى) المقبول الذي يفصل بين التصنيفات السلبية والإيجابية (Tiratira , 2009).

ويشير علام (2007) إلى أهم استخدامات الاختبارات محكية المرجع ومنها على سبيل المثال لا الحصر:

- 1- التقويم القبلي للتعرف على ما يمتلكه المتعلم من مهارات ومعارف التي تشكل الأساس للتعلم الجديد من أجل تحديد المستوى العام ونوع البرنامج والاستراتيجية المناسبة لقدرات المتعلمين.
- 2- التقويم التكويني بعد الانتهاء من الوحدة الدراسية للتعرف على نقاط القوة وأولويات التطوير والوقوف عليها من أجل التغلب عليها ومعرفة أسبابها، ومن ثم يطبق اختبار مماثل له مرة ثانية بعد فترة زمنية مناسبة.
- 3- تقويم تحصيل البرامج التدريبية أو مستوى إتقان المتعلم والذي يعبر عنه بأداء المتعلم في الاختبار الذي يطبق عليه نهاية البرنامج لتحديد مستوى الأداء المقبول.

ويمكن توضيح بعض الأمثلة على استخدامات الاختبارات محكية المرجع الشائعة:

- **اختبارات الأداء:** والتي تُقيم الأداء الفعلي للطلاب في مهارة محددة، وتشمل المشاريع العملية والعروض التقديمية.
- **الاختبارات العملية:** تستخدم في المجالات التطبيقية مثل العلوم والطب، وتتطلب من الطلاب إجراء تجارب أو تنفيذ مهام معينة.
- **الاختبارات القائمة على المشاريع:** تتضمن تنفيذ مشروع شامل يعكس المهارات والمعرفة المكتسبة، يمكن أن تشمل البحث، التصميم، أو الأعمال الفنية.
- **الاختبارات الشفوية:** تركز على قدرة الطالب على التعبير عن معرفته شفهاً، تُستخدم لتقييم الفهم العميق للموضوعات.

- الاختبارات المعتمدة على المواقف: تُقيّم قدرة الطالب على التعامل مع سيناريوهات واقعية. تُستخدم في مجالات مثل التعليم والتدريب المهني.

- التقييمات المستمرة: تشمل مجموعة من الأنشطة والمشاريع على مدار الفصل الدراسي. تساعد في قياس التقدم والتطور على المدى الطويل.

ومن الأمثلة على الاختبارات محكية المرجع في المجال التعليمي:

- الاختبارات البنائية داخل الصفوف

- الاختبارات الوطنية لمقارنة الأداء التعليمي بين المدارس.

- الاختبارات الدولية TIMSS، PISA، PIRLS

1.2. الفرق بينها وبين الاختبارات معيارية المرجع

تشير الاختبارات معيارية المرجع إلى الاختبارات المصممة لمقارنة المتقدمين للاختبار وترتيبهم بالنسبة لبعضهم البعض. وتعكس النتيجة عدد الإجابات الصحيحة التي قدمها الطالب أكثر أو أقل مقارنة بالطلبة الآخرين أو مجموعته المعيارية، وهي من الطرق شائعة الاستخدام في وضع الدرجات والتقدير للطلبة، فإذا حصل الطالب مثلاً في اختبار مادة ما على 30 من 50 فإن الدرجة 30 هي الدرجة الخام، ولن تكون لهذه الدرجة قيمة إلا عند الرجوع إلى أداء مجموعته في نفس الاختبار، ومن هنا نصل إلى الوضع النسبي للطالب بهذا المقياس.

في الاختبارات معيارية المرجع، ينظر للمادة الدراسية ككل لا كأجزاء، فامتحان اللغة العربية مثلاً يتضمن القواعد - القراءة - المحفوظات... ودرجة الطالب عليه تعني درجته في اللغة العربية كلها. ويمكن في هذه الاختبارات أن يغطي نجاح الطالب في بعض الفروع على رسوبه في فروع أخرى، فنجاح الطالب في القواعد مثلاً قد يغطي رسوبه في القراءة وبالتالي ينظر لنجاحه في مادة اللغة العربية ككل، في هذه الاختبارات تحدد في العادة درجة النجاح والدرجة العليا، فمثلاً لو كانت الدرجة الكلية (100) تكون درجة النجاح (50) وهي متوسط الدرجة والحد الأعلى (100). وذكر (علام، 2015) أن النظام مرجعي المعيار يصمم بحيث تكون التقديرات تستند إلى المنحنى الاعتيادي.

أما بالنسبة للاختبارات محكية المرجع فتقييم الطالب يكون في ضوء محكات معدة مسبقاً، وهذه المحكات تكون مطلقة وليست نسبية، فجميع الطلبة الذين يحققون نفس المستويات المحددة يحصلون على نفس التقدير، وهذا النوع يبسر على المعلم تحديد مشكلات أو صعوبات التعلم لدى طلابه، ويقترح أساليب العلاج المناسبة لذلك، لكي لا يتخلف الطالب عن أقرانه، وهذه الطريقة تكون مناسبة عندما يهتم المعلم بتمكن أو إتقان الطلبة لنطاق محدود من مهام التعلم المراد تحقيقها، وتحديد الأهداف التعليمية إجرائياً، ويمكن المقارنة بين الاختبارات محكية المرجع ومعيارية المرجع كما في الجدول التالي:

جدول (1) مقارنة بين الاختبارات محكية المرجع ومعيارية المرجع

وجه المقارنة	الاختبارات محكية المرجع	الاختبارات معيارية المرجع
الهدف	التعرف على ما استوعبه الطالب وذلك بمقارنة أداءه بمحك معد مسبقاً.	تصنيف الطلاب وفق أداء المجموعة أو ترتيب الطلبة على متصل السمة أو القدرة، فهي تنطلق من مبدأ التوزيع الطبيعي للفروق الفردية.

المقياس	يتضمن مستويات محددة من الكفاءة	يتضمن توزيع درجات الطلاب
التقييم	على خصوصيات المادة وأجزائها	على عموميات المادة ككل
استخدام البيانات	لتحديد مدى استيفاء المعايير التعليمية	لتحديد الأوائل والفئات المختلفة
النوع	من وسائل التقويم التكويني	من وسائل التقويم الختامي
المزايا	توفر تقييمًا دقيقًا لقدرة الطالب على تحقيق معايير معينة تشجع على التعلم العميق بدلاً من التنافس	سهولة التنفيذ والتصميم توفر مقارنة واضحة بين الطلاب
العيوب	قد تكون صعبة التصميم وتتطلب معايير واضحة يمكن أن تكون النتائج محدودة في تقديم صورة شاملة	قد تؤدي إلى ضغط تنافسي بين الطلاب لا تعكس دائماً مستوى معرفة الطلاب بشكل دقيق
توزيع الدرجات	توزيعات ملتوية	توزيع طبيعي
مستوى معرفة الطالب بمحتوى الاختبار	ليس للطالب فكرة عن المحتوى المتوقع أن يغطيه الاختبار	الطالب يعرف بالضبط المحتوى الذي سيغطيه الاختبار

(البريشن، 2003؛ البلوي، 2021، العزيمي، 2016)

ولا يمكن القول أن هذا النوع من الاختبارات أفضل من الآخر فلكل أهميته واستخداماته بناء على الغرض المراد من النتائج التي يتم الحصول عليها من الطلبة، فمثلاً إذا كان الغرض ترتيب مستويات الطلبة كما في الامتحانات النهائية فالاختبار المعياري هو الأنسب، بينما تصلح الاختبارات المحكية المرجع لقياس المهارات، وهكذا (البريشن، 2003).

ونظراً للاختلافات بين نوعي الاختبار (المحكي والمعيارى) فإن خصائص المفردات وثبات الاختبار في الاتجاه المحكي ينبغي أن يتم التحقق منها بطرق تختلف عن تلك في الاتجاه المعياري.

فمثلاً الاختبارات محكية المرجع لا تصمم على أساس زيادة تباين درجات المفحوصين كما هو الحال في الاختبارات معيارية المرجع والتي يتم حساب معامل الثبات فيها من خلال الصيغ التقليدية التي تتأثر بمدى تغير درجات الاختبار، بينما في الاختبارات محكية المرجع يتم تفسير الدرجة فيها بالرجوع إلى مستوى أداء محدد، فيتوقع أن تكون مجموعة الأفراد متجانسة، وخاصة عندما يكون التعليم موجهاً إلى مستوى معين من الإتقان للمعلومات أو المهارات، وبالتالي عند حساب الثبات بالطرق التقليدية للاختبارات محكية المرجع فإنه من المتوقع أن تكون القيم منخفضة، لذا فقد اقترح ليفنجستون (1970) بديلاً لثبات الاختبارات التقليدية والتي استخدم فيها درجة القطع التي بها يحدد مستوى إتقان المختبرين من عدمه (عودة، 1998؛ عابنة، 2009)، وكما اقترح بابام Popham وهاسك Husek ضم درجات مجموعتين الأولى متفنة أو تلقت التعليم والأخرى غير المتفنة أو تلك التي لم تتلقى التعليم، لتلك المهارات المراد قياسها بالاختبار، بغرض زيادة تباين درجات الاختبار، ليتمكن استخدام طرق الثبات (علام، 2007). وتوجد طرق متنوعة لحساب معامل الثبات للاختبارات محكية المرجع، ويمكن تصنيفها إلى مجموعتين: طرق تتطلب تطبيق الاختبار مرة واحدة، وطرق تتطلب تطبيق الاختبار مرتين، ولكل من هذه الطرق معادلات تختلف عن تلك المستخدمة في حساب الثبات في الاختبارات معيارية المرجع.

وبالنسبة لمفردات الاختبار محكي المرجع فهي تميل إلى السهولة نسبياً نظراً لأن المطلوب هو وصول جميع الطلبة أو أغلبهم إلى مستوى الكفاءة المناسب (المحك) أو قريباً منه، فهولا يهتم بالتمييز بين المختبرين، وإنما يقيس هدفاً مباشراً، فتحديد الصعوبة يتحدد بمستوى الهدف، فالتذكر أسهل من التحليل أو التقويم. بينما في الاختبار معياري المرجع يتم استبعاد المفردات السهلة جداً وكذلك الصعوبة جداً لأنه يسعى للكشف عن مستويات الإنجاز كافة من أدناها إلى أعلاها، فربما كان الأفضل أن تتراوح قيم الصعوبة بين (0,85-0,15) (علام، 2014).

2.2. معامل حساسية المفردة

إن الفرق الأساسي بين الاختبار محكي المرجع ومعياري المرجع هو قابلية التغير للمفردة. ويوجد نوعين من التحليل عند تحليل مفردة الاختبار محكي المرجع: التحليل المنطقي الذي يعتمد على آراء الخبراء ويتعلق بمدى مطابقة المفردة للنطاق السلوكي الذي تقيسه، والتحليل الاحصائي لدرجات المفردات والذي يتعلق بمدى جودة المفردة وتجانسها وحساسيتها لعملية التعلم (علام، 2001).

ويعرف معامل حساسية المفردة بأنه مؤشر يقيس قدرة المفردة على التفريق بين المتمكنين وغير المتمكنين في القياس المحكي، وتفتقر الكثير من الاختبارات لتحقيق هذا المؤشر بالطرق المناسبة مما يدل على وجود مشكلة في بناء هذه الاختبارات، أو هو أسلوب إحصائي يستخدم لتحليل مفردات الاختبار محكي المرجع للكشف عن المفردات التي تعكس أثر المعالجة التعليمية من خلال ما تقيسه المفردة واستجابة الأفراد لها (كرمة والحجامي، 2016؛ طيفور وآخرون، 2003).

معامل حساسية المفردة أو معامل التمييز في الاتجاه المعياري يبنى على أساس قدرته في التعرف على الفرق بين المفحوصين والذي يأخذ القيمة بين $1+$ و $1-$ لذلك هذا النوع من الاختبارات تجعل الدرجات في الاختبار تأخذ شكل التوزيع الطبيعي وبذلك يتم تحديد مستوى الفرد بناء على أداء المجموعة التي ينتمي إليها، وهو ما يتماشى مع الهدف من الاختبار المعياري، ويتم حسابه بطرق تختلف عن ما هو في الاختبارات محكية المرجع، حيث أن الهدف في الاختبارات محكية المرجع يختلف فنحن فيها نهدف للتحقق من مدى إتقان الفرد أو تمكنه من مهارة أو قدرة ما وليس معرفة مستواه بين الأفراد (Hambleton, Swaminathan & Rogers, 1999).

والمفردات الفعالة في الاختبارات محكية المرجع هي التي يجيب عليها إجابة صحيحة عدد قليل من الطلاب قبل التعليم، ويجب عليها عدد كبير منهم بعد التعليم (أبو ناهية، 1994)

وهناك أساليب متعددة للقياس الاحصائي لمفردات الاختبار محكي المرجع منها ما يعتمد على النظرية الكلاسيكية، والآخر يعتمد على نظرية الاستجابة للمفردة. والأساليب التي تعتمد على النظرية الكلاسيكية تنقسم إلى طرق تعتمد على القياس قبل وبعد التعليم، وطرق أخرى تعتمد على قياس أداء مجموعتين.

3.2. أساليب حساب معامل حساسية المفردة حسب النظرية الكلاسيكية

بالنسبة للطريقة الأولى التي تعتمد على القياس قبل وبعد التعليم فيمكن بتطبيق نفس المفردات في المرتين أو اختبارهم في المرة الثانية بمجموعة موازية أو مكافئة من المفردات، ولكن يعاب على هذه الطريقة صعوبة تطبيق الاختبار البعدي لأنه يتطلب الانتظار حتى ينتهي الطلاب من الوحدة الدراسية المطلوبة. ويمكن التغلب على هذه المشكلة عن طريق بناء صورتين متكافئتين للاختبار أو إطالة الفترة الزمنية التي تفصل بين فترتي التطبيق. ومن أمثلة الطرق التي تعتمد على القياس قبل وبعد التعلم (برينان، بابام، كوكس، فرجاس، رودابوش، كوسكوف، كلاين).

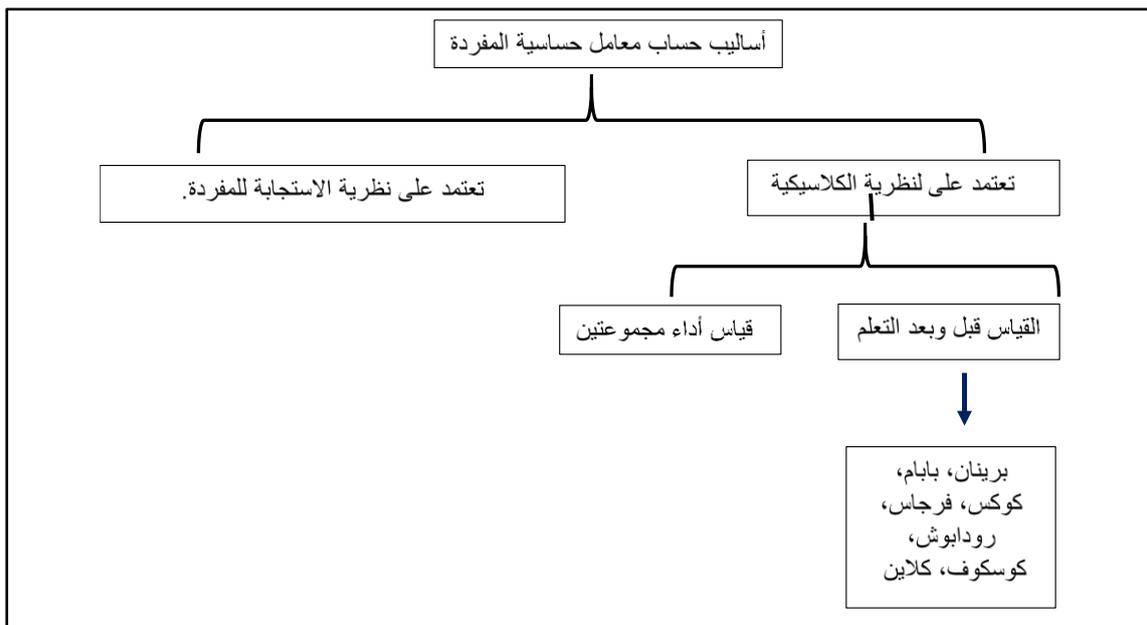
أما بالنسبة للطرق التي تعتمد على قياس أداء مجموعتين فتكون بتطبيق الاختبار على مجموعتين من الطلاب في نفس الوقت أحدهما لم تتلقى التعليم والأخرى تلقت التعليم، ويتم حساب المعامل من خلال طرح نسبة الأفراد الذين أجابوا إجابة صحيحة للمتلقين من نسبة الأفراد الذين أجابوا إجابة صحيحة لغير المتلقين ومن مميزات هذه الطرق أنه يمكن استخدامها في تحليل مفردات الاختبار في وقت واحد ولكن يعاب عليها صعوبة تعريف المحك المناسب الذي يتم في ضوءه اختيار كل من المجموعتين وكذلك التأكد من تكافؤ هاتين المجموعتين (كرمة والحجمي، 2013).

4.2. أساليب حساب معامل حساسية المفردة حسب نظرية الاستجابة للمفردة

تقوم نظرية الاستجابة للمفردة على فرضية أساسية وهي أن القيم الاحتمالية لاستجابة الفرد على مفردة تكون دالة للسمة أو القدرة التي يفترض أن يقيسها الاختبار، وتمثل العلاقة بين السمة التي يقيسها الاختبار واحتمال الاستجابة الصحيحة للمفردة ما يسمى بالمنحنى المميز للمفردة، وتكون المفردة ذات تمييز جيد عندما يكون هناك أقصى انحدار لها للسمة المراد تحقيقها، لذا فإن قدرة المفردة على التمييز بين القدرات المختلفة للأفراد يرتبط باحتمال الحصول على الاستجابة الصحيحة للمفردة بتغير هذه القدرة، أي أن احتمال الأفراد ذوي القدرة المرتفعة على المفردة إجابة صحيحة يكون أعلى من الأفراد ذوي القدرة المنخفضة، ولذلك يكون ميل المنحنى أكثر انحداراً نحو تحقيق السمة حيث أن حذف المفردات أو الإضافة لا يجعل الاختبار يتأثر بهذا الحذف أو الإضافة كما في النظرية الكلاسيكية، حيث يعتمد انتقاء المفردات من خلال تجانس مفردات الاختبار ونقصد به أن الأفراد الذين يجيبون إجابة صحيحة عن إحدى مفردات الاختبار يميلون إلى الإجابة الصحيحة عن بقية المفردات وبذلك يحصلون على درجات مرتفعة في الاختبار ويحدث العكس بالنسبة للأفراد ذوي القدرات المنخفضة لذا يطلق على معامل تجانس المفردة في نظرية IRT باسم معامل التمييز والذي يمكن الحصول عليه من خلال معامل الارتباط الثنائي المتسلسل بين درجات الأفراد على المفردة والاختبار ككل (كاظم، 1988؛ علام، 2010)

ويمكن تلخيص أساليب وطرق حساب معامل حساسية المفردة في المخطط التالي:

شكل (1) أساليب وطرق حساب معامل حساسية المفردة



(الباحثة، 2024)

3. الدراسات السابقة

هناك عدد قليل من الدراسات الأجنبية التي تناولت معامل حساسية المفردة في حدود بحث الدراسة، واختلفت نتائج هذه الدراسات ولم تتفق حول أكثر الأساليب فعالية في حساب معامل حساسية المفردة، فدراسة Haladyna (1976) أشارت إلى أن نموذج راش (أحد نماذج نظرية الاستجابة للمفردة) ومعامل الحساسية للتعلم والإحصائيات التقليدية ومعاملات بايزن متكافئة في اختيارها للمفردات، بينما أشارت دراسة Thrash (1977) إلى أن معامل كوكس-فارجاس أفضل من الطرق الأخرى لاختيار المفردات، في حين أن دراسة Silva (1985) أكدت أن النماذج المعتمدة على نظرية الاستجابة للمفردة أفضل من كوكس-فارجاس (طيفور وآخرون، 2003).

أما بالنسبة للدراسات العربية في هذا المجال فهي نادرة جدا وفي حدود البحث الحالي تم التوصل إلى دراستين فقط وهما دراسة (كرمة والحجامي، 2016) والتي هدفت لمقارنة مؤشر معامل حساسية الفقرة كدراسة محكية بين كل من طريقة كوكس وفرجاس وبين طريقة بوبام لاختبار التحليل الناقد وتبين تفوق طريقة تقدير بوبام على طريقة كوكس وفرجاس في حساب مؤشر حساسية الفقرات نحو التعليم وذلك لاعتمادها على مربع كاي في دقة تقدير الفرق وحساب معامل مؤشر حساسية الفقرة. والدراسة الأخرى دراسة (طيفور وآخرون، 2003) والتي هدفت للتعرف على أنسب طرق لحساب حساسية المفردة عند أحجام مختلفة من العينات، وعند مستويات معرفية مختلفة من الأهداف السلوكية، وكان من نتائجها أن معامل الحساسية لا يختلف باختلاف حجم العينة فقط ولكنه يختلف باختلاف مستوى الهدف السلوكي.

4. الخاتمة

إن الاختبارات محكية المرجع هي وسيلة للتحقق من مدى إتقان الطلاب للمعارف والكفايات التي تلقوها في أي برنامج تعليمي، ولها دور كبير في إدخال التحسينات في التقييم التربوي فيما لو أعدت بشكل صحيح، وأساس إعدادها الصحيح يأتي من الاختيار الدقيق للمفردات فيها، وهذا يتطلب الاعتماد على أساليب إحصائية معينة ومنها معامل حساسية المفردة. ومن خلال استقراء البحوث والدراسات في هذا المجال تبين ندرتها في الميدان العربي مما انعكس على جودة إعداد هذا النوع من الاختبارات في الدراسات العربية.

5. نتائج الدراسة:

توصلت الدراسة إلى أهم النتائج التالية:

- 1- اعتمدت مجموعة من الدراسات على تحليل المفردات في الاختبارات محكية المرجع بالاعتماد على أساليب تحليل المفردات في الاختبارات معيارية المرجع.
- 2- هناك مجموعة من الدراسات اقتصرت في تحليل المفردات على معاملات الصعوبة فقط.
- 3- ندرة الدراسات العربية التي تناولت معامل حساسية المفردة بالدراسة.

6. توصيات الدراسة:

- 1- تقديم دورات وورش تدريبية للمعلمين ولمعدي الاختبارات لتزويدهم بمهارات بناء وتحليل الاختبارات محكية المرجع.
- 2- إجراء المزيد من البحوث والدراسات في هذا المجال لإثراء المكتبة العربية من ناحية، ولتحسين إعداد الاختبارات المحكية من ناحية أخرى.

- 3- تضمين مناهج القياس والتقويم في العالم العربي لمساقات عملية في بناء وتحليل الاختبارات محكية المرجع.
- 4- توجيه القائمين ومتخذي القرار في المؤسسات التعليمية إلى أهمية الاختبارات محكية المرجع لتوفير كوادر متخصصة في هذا المجال.
- 5- إجراء دراسات لنقد مجموعة من الدراسات السابقة في بناء الاختبارات محكية المرجع وخصوصا فيما يتعلق بتحليل المفردات إحصائيا.

7. المراجع

1.7. المراجع العربية:

البريشن، خالد بن عبدالله. (2003). تفسير نتائج اختبارات التحصيل الدراسي في ضوء الاختبارات معيارية المرجع والاختبارات محكية المرجع. دراسات في المناهج وطرق التدريس، 86، 47 - 68. مسترجع من

<http://search.mandumah.com/Record/40891>

البلوي، بدر بن محمد (2021، فبراير 15). الفرق بين الاختبار محكي المرجع ومعيارى المرجع. مسترجع بتاريخ (2024 أكتوبر 15) من

https://x.com/BadrAL_Balawi/status/1361252294457704451?lang=ar/

أبو ناهية، صلاح الدين محمد. (1994). القياس التربوي. مكتبة الأنجلو المصرية.

الحموري، هند عبد المجيد والكحلوت، أحمد إسماعيل. (2009). استراتيجية مقترحة لبناء اختبارات تشخيصية محكية المرجع. مجلة العلوم التربوية والنفسية، 10(2)، 185 - 214. مسترجع من:

<http://search.mandumah.com/Record/3065>

عبانة، عماد. (2009). الاختبارات محكية المرجع فلسفتها وأسس تطويرها. دار المسيرة للنشر والتوزيع.

العريزي، عيسى بن فرج. (2016). الاختبارات التحصيلية محكية المرجع لدى معلمي المرحلة المتوسطة والثانوية: دراسة سيكومترية. مجلة كلية التربية، 61، 181 - 214. مسترجع من

<http://search.mandumah.com/Record/776595>

علام، صلاح الدين محمود (2014). الاختبارات والمقاييس التربوية والنفسية. دار الفكر العربي.

علام، صلاح الدين محمود. (2010). القياس والتقويم التربوي في العملية التدريسية. دار المسيرة للنشر والتوزيع

علام، صلاح الدين محمود. (2007). الاختبارات التشخيصية مرجعية المحك المجالات التربوية والنفسية والتدريبية. دار الفكر العربي

علام، صلاح الدين محمود. (2001). الاختبارات التشخيصية المرجعية المعيارية في المجالات التربوية والنفسية. دار الفكر العربي

عودة، أحمد. (2010). القياس والتقويم في العملية التعليمية. دار الأمل.

عودة، أحمد سليمان. (1985). القياس والتقويم في العملية التدريسية. المطبعة الوطنية.

صلاح، أحمد مراد وأمين، علي سليمان. (2012). الاختبارات والمقاييس في العلوم النفسية والتربوية. دار الكتاب الحديث. طيفور، مصطفى أحمد وشريف، نادية محمود وأبو علام، رجاء محمود وعبدالله، محمود محمد إبراهيم. (2003). طرق حساب معامل حساسية المفردة للتعلم في الاختبار المرجع إلى المحك: دراسة إحصائية مقارنة [رسالة ماجستير غير منشورة، جامعة القاهرة].

كاظم، أمينة محمد. (1988). دراسة نقدية حول القياس الموضوعي للسلوك نموذج راش. مؤسسة الكويت للتقدم العلمي. كرمة، صفاء طارق حبيب و الحجامي، بلقيس حمود كاظم. (2016). دراسة مقارنة محكية المرجع لقياس معامل مؤشر حساسية المفردات بين طريقة (COX & VARGAS) وبين طريقة (POPHAM) لاختبار التحليل الناقد. مجلة بحوث الشرق الأوسط، 70، 219-270. www.mercj.journals.ekb.eg

2.7. المراجع الأجنبية:

Haladyna, T. (1976). Comments: Measurement Issues Related to Performance

Standards. *Florida Journal of Educational Research*, 18(1), 33–34. Retrieved from <https://journals.flvc.org/fjer/article/view/133804>

Hambleton, K. Swaminathan, H. & Rogers, H. (1999). *Fundamentals of item response theory*. Sage Publication, Newbury park, CA.

Mano, C. (2017). *Beyond The Numbers A guide To Interpreting And Using The Results Of Standardized Achievement Tests*, Data Recognition Corporation, Maple, U.S.A., P. 2.

Tiratira N. L. (2009). Cut Off Scores: The Basic Angoff Method And the Item Response Theory Method. *The International Journal of Educational and Psychological Assessment*, 1 (1), 27-35.

Doi: <https://doi.org/10.52133/ijrsp.v6.62.4>