بناء اختبار محكي المرجع في مادة الرياضيات للصف السابع ودراسة خصائصه السيكومترية

ريما اليوسف*

الملخص

هدفت الدراسة الحالية إلى بناء اختبار محكي المرجع في مادة الرياضيات للصف السابع، تم بناء الاختبار المحكي المرجع وتكوّن من (59) بنداً ، وطبق الاختبار على عينة مكونة من (580) طالباً وطالبةً من الصف السابع، وشمل الاختبار المنهاج كاملاً واستخدم المنهج الوصفي التحليلي، جرى دراسة صدق المحكمين والصدق الداخلي تم بعد تطبيق الاختبار والمكوّن من (77) بنداً بصورته الأولية على عينة استطلاعية، وكذلك تمت دراسة الصدق التمييزي وكان الاختبار صادق تمييزياً. كما تم دراسة الثبات باستخدام معامل كورد ريتشارد سون–21 والثبات بالتجزئة النصفية والثبات بالإعادة وجاءت القيم مرتفعة، كما تمت دراسة سهولة البنود وقدرتها على التمييز تم استبعاد البنود التي كانت صعبة أو معامل تمييزها منخفض وبلغ عددها (18) بنداً ، ثمّ تم التطبيق على عينة الدراسة وتحليل إجاباتهم **وتوصلت الدراسة الى:** بلغ معامل الصدق التلازمي والتمييزي للاختبار (0.841) وهو معامل ارتباط مرتفع ، كما حُسب الثبات بطرق متعددة وجاءت القيم عالية، كما تم إعادة تطبيق الاختبار وجاءت النتائج مرتفعة بلغت (0.88)، وكذلك تم إجراء التحليل واستخراج معاملات السهولة والتمييز حيث جاءت ضمن الحدود الطبيعية ماعدا مجموعة من البنود تم حذفها.

الكلمات المفتاحية: الاختبار المحكي المرجع.

*بحث تسجيل دكتوراه، قسم القياس والتقويم التربوي والنفسي، كلية التربية، جامعة دمشق..
Rema.Alyosef@hotmail

construction of Criterion - Referenced test in mathematics for the seventh grad

Rema Al Yousef *

Abstract

The current research examines the use of classical theory in the construction of the test of reference in the mathematics of the seventh grade, where the Criterion – Referenced test was constructed of (50) single choice of multiple under four alternatives, one of which represents the correct answer.

The test was carried out on a sample of 580 students from the seventh grade. The test included the entire curriculum and the analytical descriptive method was applied at the end of the academic year.

The accuracy of the correlation coefficients was statistically significant,with the highest value of 0.96.

The validity of the test was also verified. Stability was also studied using the Cord-Sung-21 coefficient, as well as semi-retail stability and repeatability (high). As for the ease of vocabulary, it was calculated for the test items, and they were all within the normal limits except items (1-22-29) and therefore excluded. Also, for the ability indicators, the distinguishing characteristics of the test were within the natural limits except for (18)Test items (59) items.

The Wright-Smith map shows that individuals tend to be centered around their average power value within two standard deviations except for a small number Of the individuals did not exceed (12) individuals, and the stability factor was studied as well .Thus, the study achieved its objectives, the most important of which is to obtain a test included in mathematics for the seventh grade. As for the ability of discrimination, the values of the coefficients of discrimination of items were consistent between them and restricted between (1.32 - 0.79) Exceed (0.10) and the largest (0.06).

Key word: classical theory ،criterion referenced test

* Phd registration – doctoral degree ,Evaluation and Measurement Department.**Rema.Alyosef@hotmail.**

المقدمة:

يحتل القياس مكانة هامة في العلوم المختلفة، ويعد سمة هامة من سمات التقدم العلمي في المجالات المختلفة، والقياس في التربية يسعى إلى الكشف عن حقيقة الظاهرة التي يتصدى لها كما يسعى إلى التحديد الكمي لهذه الظاهرة، فمنذ أن وجدت حركة القياس النفسي والتربوي اهتم علماء النفس بتحقيق صدق الاختبارات والمقاييس وثباتها سعياً منهم لتحقيق أعلى درجة من الموضوعية في هذه الأدوات عند استخدامها في عملية القياس.(عودة، 2010، 56).

تعد الاختبارات إحدى أهم أدوات القياس فهي من أهم الوسائل وأكثرها فاعلية لقياس التحصيل الدراسي لدى الطالب حيث يمكن من خلالها معرفة مدى التقدم في تحقيق الأهداف المطلوبة وتوفير معلومات وبيانات للتغذية الراجعة المستمرة وفي عام(1963) دخل قاموس المصطلحات في مجال القياس والتقويم مصطلح جديد وهو الاختبارات محكية المرجع. (مخائيل، 2010، 277).

يرى علام أن الاختبارات محكية المرجع تعد أكثر أنواع الاختبارات التحصيلية ملائمة لقياس تحصيل الطالب وتقويمه لأنها تقوم بتحديد المهارات والكفايات المطلوب إتقانها بدقة فائقة تمكن المدرس من قياسها وملاحظتها بشكل مباشر ومن ثم تقدير ما حققه الطالب من تلك المهارات والكفايات بناءً على مستوى أداء محدد.(علام، 2001، 43).

وتعد مادة الرياضيات مادة أساسية تدرس في جميع المراحل لا يمكن التغاضي عن مدى أهميتها ودورها في الحياة، فهي علم متتابع ومتكامل، وتعتبر الرياضيات علماً مجرداً ومنظماً يصل إلى أي نتيجة من خلال عرض وتفسير وتحليل البيانات والأفكار.(سنقر،2000، 130).

مشكلة البحث:

بحث الكثير من العلماء في مفهوم الاختبارات التحصيلية، ولعل أبرز التوجهات في تحديد هذا المفهوم هو ربطه بمفهوم التعليم الدراسي، فقد استخدمت اختبارات التحصيل لتحديد ما تعلمه الفرد بعد أن تلقى نوعاً معيناً من التعليم (رضوان، 2011، 56).

وتعد الاختبارات التحصيلية من أهم وسائل تقويم التحصيل، وتحديد مستوى المتعلمين التحصيلي وقد أشارت الدراسات والبحوث التي أجريت على الاختبارات التي يعدها المعلمون إلى وجود نقاط ضعف عدة فيها. لذلك لابد أن يوليها المعلم أهمية كبرى في التخطيط والإعداد (أبو سل، 2002، 55).

لذلك حرص العاملون في الميدان التربوي وغيرهم من التربويين على أن تكون هذه الاختبارات ذات كفاءة عالية في عملية القياس والتقويم وهذه الكفاءة لا تتأتى إلا من خلال إعداد اختبارات نموذجية وفاعلة تخلو من الملاحظات(علام، 2001، 25).

لذا نشأت فكرة الاختبارات محكية المرجع نتيجة رغبة المتخصصين في تعرف وسائل جديدة تمكنهم من الوصول إلى معلومات أكثر عن مستوى أداء الطلبة، وتعنى هذه الاختبارات بدرجة الكفاية في مهارات محددة أي أن التركيز فيها يكون على مدى وصول الفرد إلى مستوى الأداء في مهارة ما، فهي ترتكز على مجموعة من الأهداف الأكثر تحديداً وتفصيلاً وتبعاً لذلك فإنه يتم تصميمها بحيث تكون على درجة عالية من الموافقة وصدق المحتوى بفضل ارتكازها مباشرة على الأهداف (ميخائيل، 2003، 297).

وكثيراً ما تستخدم الاختبارات محكية المرجع في استراتيجيات التدريس والتعلم فيها للتمكن من الأداء، كما أنها لا تهتم بتحديد مكانة الفرد بين أقرانه وبالتالي لا تستبعد أياً من البنود مادامت تقيس هدفاً سلوكياً مهماً (علام، 2005، 427).أما بالنسبة لمادة الرياضيات فهي من العلوم الهامة التي تؤثر في طريقة التفكير لدى الإنسان فتجعله منظماً كما أنها تنمي مهارات الإنسان الحياتية وطرق التواصل وطريقة توليد الأفكار

الجديدة ، لذلك نحن بحاجة إلى بناء اختبارات ذات خصائص سيكومترية جيدة تساعد المعلمين في قياس تحصيل الطلبة.

وبالتالي تتلخص مشكلة البحث في السؤال الآتي:

ما هي الخصائص السيكومترية للاختبار المحكي المرجع في مادة الرياضيات للصف السابع؟

ثانياً: أهمية البحث:

1-أهمية الاختبارات التحصيلية عامة وانتشارها الكبير واستخدامها في المؤسسات التعليمية والاعتماد عليها في قياس تحصيل الطلبة.

2-أهمية الاختبارات المحكية المرجع التي تعزز ثقة الفرد بقدراته بدلاً من مقارنة أدائه بأداء مجموعته.

3- بناء اختبار محكي المرجع لمادة الرياضيات يساعد المعلمين في استخدامه لقياس تحصيل الطلاب في مادة الرياضيات.

4-بناء أداة قياس ذات خصائص سيكومترية جيدة يمكن أن تستخدم لقياس المهارات المتطلبة في مادة الرياضيات.

5-قياس قدرات الطلاب وتشخيص جوانب القوة والضعف.

6-أهمية المرحلة الدراسية التي يتناولها البحث وهي طلاب الصف السابع من التعليم الأساسي حيث يجب بناء اختبارات وفق القواعد التي تلائم طبيعة القدرات العقلية

7- الوصول إلى اختبارات بدرجة عالية من الجودة يتطلب استخدام أكثر من طريقة لتحديد فقرات الاختبار.

ثالثاً: أهداف البحث:

يهدف البحث إلى:

1-بناء اختبار محكي المرجع في مادة الرياضيات للصف السابع.

2-دراسة صدق الاختبار المحكي المرجع في مادة الرياضيات للصف السابع.

3-دراسة ثبات الاختبار المحكي المرجع في مادة الرياضيات للصف السابع.

4-دراسة القدرة التمييزية للاختبار المحكي المرجع في مادة الرياضيات للصف السابع .

5-دراسة صعوبة البنود للاختبار المحكي المرجع في مادة الرياضيات للصف السابع.

رابعاً: أسئلة البحث:

السؤال الأول: ما مؤشرات صدق الاختبار المحكي المرجع في مادة الرياضيات؟

السؤال الثاني: ما مؤشرات ثبات الاختبار المحكي المرجع في مادة الرياضيات؟

السؤال الثالث: ما مؤشرات القدرة التمييزية للاختبار المحكي المرجع في مادة الرياضيات؟

السؤال الرابع: ما مؤشرات صعوبة الفقرات للاختبار المحكي المرجع في مادة الرياضيات؟

خامساً: منهج البحث:

تم اعتماد المنهج الوصفي التحليلي لبناء أداة الدراسة وتفسير النتائج ويهتم المنهج الوصفي التحليلي بدراسة الظاهرة ووصفها وصفاً دقيقاً ويعبر عنها تعبيراً كمياً أو كيفياً.(فرج، 2007، 95).

سادساً: مجتمع البحث وعينته:

تكون من تلاميذ الصف السابع جميعهم الذين يدرسون مقرر الرياضيات في عينة من المدارس الرسمية في محافظة دمشق للعام الدراسي 2018/2017م والبالغ عددهم /26036/ طالباً وطالبة، وقد بلغ عدد الذكور /12783/ وعدد الإناث /13253/.

أما عينة البحث فتكونت عينة البحث من (580) تلميذاً وتلميذة من الصف السابع وقد تم اختيارهم بطريقة عشوائية، من مدارس محافظة دمشق بطريقة عشوائية، وبالنسبة للعينة الاستطلاعية تكونت من (100) تلميذاً وتلميذة من الصف السابع من التعليم الأساسية وجرى اختيارهم من إحدى المدارس في مدينة دمشق من خارج عينة الدراسة.

سابعاً: حدود البحث:

الحدود المكانية: تمّ تطبيق الدراسة في المدارس الرسمية في محافظة دمشق.

الحدود البشرية: تلاميذ الحلقة الثانية من التعليم الأساسي (الصف السابع).

الحدود الزمانية: جري تطبيق الدراسة في العام الدراسي 2018م/2019م.

ثامناً: أداة البحث:

- اختبار محكي المرجع لمقرر الرياضيات للصف السابع بني استناداً إلى المعايير المعتمدة من قبل وزارة التربية لكتاب الرياضيات.
- برنامج (Excel) لحساب معاملات السهولة والتمييز.
- حزمة البرامج الإحصائية (spss) لحساب الصدق والثبات.

تاسعاً: إجراءات البحث:

- بناء اختبار محكي المرجع لمقرر الرياضيات للصف السابع وفق عدة خطوات:
- تحليل محتوى المقرر الدراسي في ضوء الأهداف المحددة.
- صياغة أسئلة مناسبة للأهداف الدراسية المحددة في المقرر بالاستناد إلى المعايير المعتمدة لدى وزارة التربية لكتاب الرياضيات للصف السابع.
- تحديد شكل بنود الاختبار وهي بنود اختيار من متعدد بأربعة بدائل
- وضع قائمة من البنود المقابلة للمعايير والمؤشرات المعتمدة من قبل وزارة التربية
- عرض البنود على المحكمين والأخذ بملاحظاتهم وقد بلغ عدد البنود سبعة وسبعين بنداً.

- تجريب الاختبار على عينة استطلاعية من خارج عينة البحث الأساسية للتعرف على مدى وضوح التعليمات والأسئلة والزمن اللازم للإجابة ودراسة معاملات السهولة والتمييز لبنوده وصدقه وثباته حيث بلغ عدد أفرادها (100) تلميذاً وتلميذة، حيث تم دراسة صدق المحتوى من خلال عرضه على مجموعة من المحكمين والاستفادة من الملاحظات وإجراء التعديلات المطلوبة ، كما تم دراسة الصدق البنيوي من خلال حساب معامل الارتباط بيرسون بين الدرجة الكلية وعلامة الفرد على كل بند للتأكد من الصدق البنيوي حيث جاءت القيم كلها دالة إحصائياً عند مستوى دلالة أقل من /0.01/ وكانت أقل قيمة /0.430/ وأعلى قيمة /0.965/.

أما بالنسبة للصدق التمييزي تم بعد تطبيق اختبار t-test وأشارت النتائج إلى وجود فروق ذات دلالة إحصائية بين متوسطي درجات المجموعتين العليا والدنيا ما يدل على قدرة الاختبار على التمييز بين المستويات المختلفة للطلاب، أما بالنسبة للثبات فقد تم حسابه من خلال الاتساق الداخلي باستخدام معامل كودر –ريتشاردسون21 حيث بلغت /0.87/ وهي قيمة ثبات عالية والاختبار يتمتع ببنية داخلية ثابتة، كما تم حساب معاملات السهولة والتمييز لمعرفة مدى سهولة بند ما في الاختبار ومدى قدرته على التمييز بين الطالب ذي القدرة العالية والطالب ذي القدرة المنخفضة ،حيث كانت المعاملات ضمن الحدود المطلوبة ماعدا مجموعة من البنود تم حذفها وبعد حذف البنود الغير مناسبة أصبح الاختبار مكون من (59) بنداً.

- إجراء التعديلات وفقاً لنتائج الدراسة الاستطلاعية.
- إخراج الاختبار بشكله النهائي.
- تطبيق الاختبار محكي المرجع لمقرر الرياضيات على عينة الدراسة المكونة من (580) تلميذاً وتلميذة من تلاميذ الصف السابع في مدارس محافظة دمشق.
- تقدير درجات الطلبة على الاختبار ،وأعطيت درجة واحدة للإجابة الصحيحة وصفر للإجابة المغلوطة.

- تحليل البيانات واستخراج النتائج.
- التحقق من صدق الاختبار والثبات واستخراج معاملات السهولة والتمييز باستخدام حزمة البرامج الإحصائية (spss) وبرنامج (Excel).

عاشراً: المصطلحات والتعريفات الإجرائية:

- **الاختبار المحكي المرجع:** هو الاختبار الذي يتم فيه مقارنة أداء الفرد بمستوى أداء محدد مسبقاً دون مقارنته بمتوسط أداء جماعة مرجعية(علام،2001،26).

وتعرفه الباحثة: هو الاختبار الذي تم إعداده والذي لقياس درجة إتقان تلاميذ الصف السابع/ الحلقة الثانية من التعليم الأساسي لمادة الرياضيات.

- **الخصائص السيكومترية:** دلائل أو مؤشرات إحصائية على مدى جودة المقياس وفقراته إذ توجد خصائص سيكومترية للفقرات وهي تمييز الفقرة واتساقها الداخلي أي صدقها وتوجد خصائص سيكومترية للمقياس وهي صدقه وثباته وحساسيته وشكل التوزيع التكراري للدرجات(مرجع سابق،207).

وتعرفها الباحثة: هي دراسة صدق وثبات الاختبار وكذلك معاملات السهولة والصعوبة والتمييز.

- **الرياضيات:** لم يتفق جميع العلماء على تعريف موحد للرياضيات، فقد تغير المفهوم وتطور عبر العصور والأزمنة، إذ بدأ بسيطاً حيث بحث في عمليات الجمع والطرح الأساسية وتطور شيئاً فشيئاً حتى أصبح علماً يبحث في أكثر الأمور تعقيداً، فالرياضيات هو العلم الذي يبحث في عدة أمور مختلفة كالكم والقياس والحساب والهندسة ووصف الأشكال والعد ويتم باستخدام الأرقام والرموز (راشد، 2005، 13).
- **طلاب الصف السابع:** هم طلاب الحلقة الثانية من التعليم الأساسي وتشمل من الصف السابع إلى الصف التاسع.

- **الجبر:** أحد فروع علم الرياضيات، ويقوم على مبدأ المعادلات التي تتكون من مجموعة من المتغيرات والثوابت بحيث تكون الثوابت عبارة أرقام ذات قيمة محددة لاتتغير.
- **الهندسة:** هي فرع من فروع الرياضيات المعنية بدراسة الأشكال، وقياس الحجوم والمساحات، ودراسة الهندسة الفراغية.
- **الاحتمال:** تحليل الحوادث العشوائية، فمن غير الممكن معرفة نتائجها الحتمية قبل حدوثها، ولكن معرفة النتائج المحتملة لها من الممكن أن تجعل التنبؤ بالنتيجة الفعلية ممكناً، وتعد التجربة التي يمكن تكرارها عملياً أو افتراضياً أهم عنصر لدراسة الاحتمالات.
- **القياس:** هو طريقة لوصف الأحداث أو الأجسام أو غيرها عن طريق الأرقام، يمكن من خلاله حساب المساحات للأشكال المختلفة وقياس المساحات.
- **معالجة البيانات والتفسير البياني :** قراءة وتفسير مخطط الانتشار وحل المسائل بالاعتماد على البيانات الممثلة بمخطط الانتشار
- **الصدق:** يعرف بأنه المدى الذي يحقق به الاختبار الغرض الذي وضع من أجله، وتعتبر صفة الصدق مهمة للاختبار الجيد ويقصد به أن يقيس الاختبار ما صمم لقياسه.
- **الثبات:** من العوامل الهامة والواجب توافرها لصلاحية استخدام الاختبار ويتعلق بدقة القياس بغض النظر عما يقاس ويعزى الثبات إلى النتائج وليس إلى الاختبار نفسه.
- **معاملات الصعوبة:** نسبة الأفراد الذين أجابوا عن كل بند إجابة صحيحة ، وكلما زادت هذه النسبة دل على سهولة البند وكلما قلت دل على صعوبة البند.
- **معاملا السهولة :** كلما ارتفعت قيمة معامل الصعوبة كلما اعتبر البند سهلاً، بمعنى أن مؤشر السهولة يشير إلى نسبة الأفراد الذين أجابوا إجابة خاطئة على البند.

- **معاملات التمييز:** مدى قياس بنود الاختبار للفروق الفردية بين الأفراد، فقدرة البند على التمييز هو قدرته على التفريق بين الأفراد الأكثر مهارة في قدرة معينة والأفراد الأقل مهارة في نفس القدرة.
- **تعريف درجة القطع** : هي الدرجة التي تفصل بين الطلاب المتمكنين والطلاب الغير متمكنين .

حادي عشر: الإطار النظري للبحث:

أولاً: مفهوم القياس:

يشير القياس إلى الجوانب الكمية التي تصف خاصية معينة، أما في المجال التربوي والنفسي فالقياس له أهمية كبيرة في تعيين رقم معين لعملية معينة، كأن يقاس تحصيل الطلبة من خلال اختبار تحصيلي فهنا يقدم القياس درجة أو علاقة محددة للطالب تدل على مقدار ما حصلته من معلومات (علام، 2006، ص15).

يعد القياس أمراً هاماً في مجال التربية وعلم النفس، إذ من خلاله يمكن قياس خاصية معينة، والنتيجة عبارة عن أرقام أو رموز تناظر الخاصة المقاسة وهذه الأرقام يتم التعبير عنها من خلال قيم عددية أو درجات تعبر عن العلاقة القائمة بين مستويات الخاصية المقاسة لدى الأفراد. (Thorndik, et al, 1991, pp 9).

ثانياً: الاختبارات المحكية المرجع:

- **تعريف الاختبارات المحكية المرجع:**

يرجع الفضل في ظهور مصطلح الاختبارات محكية المرجع إلى روبرت جلازر والذي قدمه في مقال في مجلة علم النفس الأمريكية، إذ توقع أنصارها الأوائل بأن هذه الاستراتيجية سوف تحدث ثورة في التعليم (Popham,2014,62).

يطلق على هذه الاختبارات تسميات متعددة منها اختبارات التفوق و اختبارات الإتقان واختبارات المهارات الأساسية، فقد ميز جلازر في أحد دراساته بين الاختبار محكي المرجع والاختبار معياري المرجع على أساس أداء الفرد في الاختبار محكي المرجع

ينسب إلى محتوى الاختبار ذاته وليس إلى أداء الأفراد الآخرين كما في الاختبارات معيارية المرجع (كاظم، 1988، ص64).

ويكون التركيز في هذه الاختبارات على مدى وصول الفرد إلى مستوى الأداء في مهارة ما، لذا يقاس أداء الفرد بالمستوى الذي يحدد الفاحص سابقا فإذا تمكن هذا الفرد من الوصول إلى المهارات المطلوبة سابقاً فهذا يعني أنه يمكن أن يجتاز المرحلة التي وضع لها هذا الاختبار (مرجع سابق، ص85) وأشار جلازر أن الاختبار محكي المرجع هو نوع من أنواع الاختبارات المصمم لقياس تحصيل الفرد استناداً إلى محك سلوكي نوعي للكفاءة .(ميخائيل، 2012، 278).

في حين يرى براين (Brian) أن الاختبار المحكي المرجع هو ذلك الاختبار الذي يتناول موضوعات وكفايات محددة يتم مقارنتها في ضوء معيار محدد وهو ما نسميه درجة القطع. (Brian,2009,2).

- **صفات الاختبارات المحكية المرجع:**

– ترتكز على مجموعة من الأهداف السلوكية أو الإجرائية الأكثر تفصيلاً وتحديداً

_ يتم تصميمها بحيث تكون على درجة عالية من صدق المحتوى بفضل ارتكازها مباشرة على الأهداف.

– يفسر أداؤها استناداً إلى محك معين يحدد مسبقاً ويمثل الحد الأدنى المقبول للأداء

– الطالب إما ناجح أو يعيد حتى يصل إلى مستوى الإتقان.

– لا تهتم بتحديد مكانة الفرد بين أقرانه.

- **درجة القطع:**

تتعدد آراء علماء القياس والتقويم التربوي في تسميات درجة القطع للاختبارات المحكية المرجع ومن تلك المسميات:

- درجة القطع.
- درجة النجاح.
- الحد الأدنى من الكفاية.

- مستوى التمكن. (عبابنة، 2009، ص127).

- **طرائق تحديد درجة القطع:**

توجد طرائق متعددة لتحديد درجة القطع في الاختبارات المحكية المرجع إذ ذكر هامبيلتون أن هماك من 10 إلى 20 طريقة رئيسية لتحديد درجة القطع ويوجد بينها نقاط اختلاف كبيرة وتختلف هذه الطرائق من حيث أسلوب (Hambelton,2009,45) بنائها وتطبيقها وتفسير نتائجها وهذه الطرائق هي:

- **الطرائق التحكيمية:** تعتمد هذه الطريقة على آراء المختصين في مجال القياس أو المختصين في محتوى الاختبار المراد قياسه، حيث يستعين بآرائهم في تحديد مستوى المتعلم من خلال تحديد عدد الأسئلة التي يجب أن يجيب عنها بشكل صحيح.

وتتفرع عن هذه الطريقة طرائق متعددة نذكر منها:

- **طريقة أنجوف:** حيث يفحص كل محكم كل مفردة من مفردات الاختبار ثم يتصوروا مجموعة من الأفراد الذين حققوا الحد الأدنى للكفاية التي يقيسها الاختبار ثم يقدر نسبة الأفراد الذين يحتمل أن يجيبوا إجابة صحيحة عن كل مفردة من مفرداته، ومتوسط هذه النسب هو الحد الأدنى لمستوى الاجتياز في الاختبار.
- **الطرائق التي تعتمد جزئياً على التحكيم وتسترشد ببيانات تجريبية:** اقترح علماء القياس العديد من الطرائق المختلفة لتحديد مستويات الأداء التي تتضمن طرائق تجريبية والحصول على مستويات أداء دقيقة يمكن الاعتماد عليها.

ومن هذه الطرائق:

- **الطرائق التحكيمية المعززة بالمعلومات:** تعتمد على بيانات مسبقة عن أداء المتعلمين توضع في حساب المحكمين لكي يسترشدوا بها وتتضمن إحصاءات خاصة بالبنود وبعض تعليقات الخبراء وأحكامهم عن البنود واقتراح المختصين وتقدم جميع هذه المعلومات للمحكمين لفحصها ومناقشتها وإعادة النظر فيها.

- **طريقة أنجوف المعدلة:** تزويد المحكمين ببعض البيانات المستمدة من التجريب الميداني إذ يطلب من كل محكم تقدير احتمال لإجابة الأفراد ذوي الحد الأدنى من الكفاية المطلوبة لبنود الاختبار ككل وعند الحصول على هذه التقديرات يتم تزويد الحكمين ببيانات تجريبية عن بنود الاختبار تشتمل على متوسط الصعوبة لكل بند ثم يطلب منهم تعديل تقديراتهم في ضوء تلك البيانات التجريبية ويكون الحد الأدنى لمستوى الاختبار هو متوسط مجموع تقديراتهما المعدلة (علام،85،2001).

وبالنسبة للاختبار الحالي تم تحديد درجة القطع بالاعتماد على الطريقة التي تعتمد على تقديرات المحكمين وآرائهم وهي طريقة أنجوف.

- **خطوات بناء الاختبارات المحكية المرجع:**

يتطلب الاختبار المحك المرجع اتباع مجموعة من الخطوات ويمكن إجمالها فيما يلي:

1. تحديد النطاق السلوكي الذي يقيسه الاختبار وتحليله إلى مكوناته من معارف ومهارات .
2- إعداد قائمة مفصلة بالأهداف التعليمية قبل البدء بعملية التعليم والتقويم وبحيث تتم صياغة هذه الأهداف بالدرجة القصوى من التحديد والتخصيص
3- اتخاذ قرار حول محك للأداء يدل على تحقق الأهداف التعليمية ووصول المتعلم للمستوى المرغوب.
4- تحديد الشكل الذي ستأخذه البنود الاختبارية وتطوير عدد من البنود لكل هدف وبحيث تكون هذه البنود عينة صادقة وممثلة للهدف المعطى.

5– المراجعة والتحليل المنطقي للبنود الاختبارية من قبل واضع الاختبار نفسه ومجموعة من المعلمين والمختصين بهدف الكشف عن العيوب المحتملة فيها، غير أن عملية المراجعة والتحليل المنطقي وحدها لا تكفي ولابد أن تتبعها عملية هامة وهي التطبيق التجريبي للبنود حيث يمكن من خلاله تفحص إجابات الطلبة قبل التعليم وبعده لتحديد ما إذا فشلت البنود في الكشف عما أعدت لكشفه.

6– إخراج الاختبار في صورته النهائية بعد أن تمت تصفيته كما تم التأكد من تمثيل كل مجموعة من البنود التي يضمها للهدف الذي أعدت له .(ميخائيل، 2010، ص154).

3– الصدق والثبات:

الصدق: يعد شرطاً أساسياً من شروط القياس ويقصد بالصدق أن يقيس الاختبار ما وضع لقياسه ويعرفه ليندكوست Lindquist أن الصدق هو الدرجة التي يقيس بها الاختبار ما يراد قياسه. (فرج، 2007، 240).

– أشكال الصدق:

– صدق المحتوى (content validity):

يشير إلى مدى تطابق فقرات الاختبار مع مضمون الاختبار أو محتواه أو هدفه، ويكون الاختبار صادقاً إذا كانت فقرات الاختبار تمثل المجتمع الأصلي للاختبار ويستخدم صدق المحتوى في اختبارات التحصيل لقياس الدرجة التي أتقن بها المتعلم مقرراً معيناً. (مراد وسليمان، 2002، 351).

– الصدق المحكي:

الدرجة التي يرتبط عندها الأداء على الاختبار بالأداء على اختبار آخر عد محكاً للاختبار الأول وأساساً في الحكم على صلاحيته.(ميخائيل، 2010، 84).

ويوجد أنواع للصدق المرتبط بالمحك وهي:

- **صدق تنبؤي predictive validity**: قدرة الاختبار على التنبؤ بأداء الفرد في موقف مستقبلي إذا كان هذا الموقف له علاقة بما يقيسه الاختبار.
- **صدق تلازمي**: يقصد به حجم العلاقة بين درجات المفحوصين على الاختبار ودرجاتهم على اختبار آخر (محك) والذي يتم الحصول عليها في الوقت نفسه فهو يهدف للتعرف إلى أداء الأفراد وتشخيصه في الوضع الراهن. (علام، 2000، 209).

أنواع المحكات: يوجد محكات متعددة نذكر منها : محك التحصيل الدراسي، محك الأداء في برنامج تعليمي، طريقة الفرق المتقابلة، طريقة المجموعات الطرفية.

وتعرف طريقة المجموعات الطرفية: قدرة الاختبار على التمييز بين طرفي القدرة التي يقيسها وتعتمد على المقارنة وحساب دلالة الفروق، غير أن المقارنة هنا تخص الفئتين العليا والدنيا من أفراد مجموعة واحدة ولا تخص مجموعتين من الأفراد.

ويمكن إجراء المقارنة بإحدى الطريقتين التاليتين:

- مقارنة الفئات المتطرفة في الاختبار والمحك الخارجي.
- مقارنة الفئات المتطرفة في الاختبار نفسه وكثيراً ما تستخدم هذه الطريقة عند دراسة القدرة التمييزية للاختبار في مراحل الإعداد والبناء(Anastasi,1982,p.142).

- **الصدق البنيوي construct validity:**

يشير إلى الدرجة التي يقيس معها الاختبار السمة أو الخاصية التي يفترض أن يقيسها. (عودة، 1993، 384).

والصدق البنيوي مفهوم شامل يتضمن الأنواع الأخرى من الصدق ويتطلب اللجوء إلى الاستنتاجات المنطقية بالإضافة إلى الوسائل التجريبية والإحصائية.

توجد طرائق متعددة للتحقق من الصدق البنيوي نذكر منها:

1– الطرائق التي تعتمد على التحليل المنطقي: ترتكز على الفحص الدقيق لبنود الاختبار والأداء الذي تتطلبه ومقابلة نتائج هذا الفحص بالنظرية المعتمدة في بناء الاختبار والتفسيرات التي قد تنبع عنها.

2– **الترابط مع اختبارات أخرى:** يتوقع أن تترابط درجات الاختبار التجريبي ترابطاً موجباً مع درجات اختبار آخر يفترض أنه يقيس السمة نفسها وأن تترابط بدرجة أقل مع درجات اختبارات تقيس سمة أخرى.

3– **التحليل العاملي:** يؤدي دوراً بالغ الأهمية في تقدير الصدق البنيوي من خلال الكشف عن السمة التي يقيسها الاختبار بأسلوب كمي إحصائي، فالهدف الرئيس له هو تبسيط عملية وصف السلوك عن طريق خفض عدد المتغيرات التي تقيسها الاختبارات إلى عدد ضئيل من العوامل(مخائيل، 2005، 172).

– الثبات Reliability:

يقصد به الدقة والإتقان الذي يقيس به الاختبار الظاهرة أو السمة التي وضع من أجلها، ويجب أن نحصل على نتائج لها صفة الاستقرار عند تكرار تطبيق الاختبار، كما أنه يجب ألا يتغير وضع المفحوص بالنسبة لمجموعته إذا أعيد تطبيق الاختبار تحت نفس الظروف.

1– طرائق حساب الثبات:

– طريقة الثبات بالإعادة test- retest method:

وهي تطبيق الاختبار على مجموعة من الأفراد ثم يترك هؤلاء الأفراد فترة من الزمن ويعاد تطبيق الاختبار نفسه على الأفراد أنفسهم وتحت الظروف نفسها ثم يحسب معامل الارتباط بين التطبيقين.

- **طريقة التجزئة النصفية split-half reliability:**

يتم تقسيم الاختبار بعد تطبيقه في جلسة اختبارية واحدة إلى نصفين يفترض أنهما متكافئان، ثم حساب معامل الارتباط بين الدرجات التي حصل عليها المفحوصين على النصف الأول والدرجات التي حصلوا عليها على النصف الثاني. (ميخائيل، 2010، 195).

- **طريقة كودر– ريتشاردسون kuder-Richardson:**

وهي تجزئة الاختبار الواحد إلى عدد كبير من الأجزاء يتكون كل جزء من هذه الأجزاء من بند واحد فقط من بنود الاختبار وكلما كان هناك اتساق بين هذه البنود كان ثبات الاختبار أعلى.

- **طريقة الصور المتكافئة Equivalent forms method**

يتم إعداد صورتين متكافئتين للاختبار وتطبيقهما على نفس الأفراد ويجب أن تكون متكافئتين في عدد البنود ونوعها والزمن المخصص للإجابة وطريقة الإجابة والتصحيح والتعليمات وتساوي معاملات الصعوبة والتمييز والانحراف المعياري والمتوسط الحسابي. (مراد وسليمان، 2002، 361).

- **طريقة معامل ألفا coefficient Alpha method:**

وهي مشتقة من صيغة عامة من معادلة كودر–ريتشاردسون20 وتختصر بمعادلة ألفا.

- **طريقة التحليل العاملي:** بالإضافة للطرق السابقة في دراسة الثبات هناك طرق أخرى تجدر الإشارة إليها، لعل من أهمها التحليل العاملي الذي يستهدف الوصول إلى أقل عدد من العوامل التي يمكن أن تفسر الأداء على الاختبار استناداً إلى دراسة الترابطات بين البنود التي يتكون منها الاختبار والكشف عن تلك التجمعات من البنود التي يتشبع كل منها بعامل معين دون غيره من العوامل (ميخائيل،2005،205).

ثاني عشر: بعض الدراسات السابقة:

- **الدراسات العربية:**

1- دراسة العنزي (2004):

عنوان الدراسة: بناء اختبار تحصيلي لقياس درجة إتقان المهارات الأساسية في مادة الرياضيات لدى تلاميذ الصفوف العليا في المرحلة الابتدائية

أهداف الدراسة: بناء اختبار تحصيلي في مادة الرياضيات يتم بواسطته تحديد المهارات الأساسية في هذه المادة للصفوف العليا (رابع، خامس، سادس) ومعرفة درجة إتقان تلاميذ الصفوف العليا هذه المهارات

عينة الدراسة: طلبة الصفوف العليا من المرحلة الابتدائية (رابع، خامس، سادس) في مدارس مدينة مكة المكرمة

أداة الدراسة: تم بناء ثلاثة اختبارات محكية المرجع لكل صف دراسي على حده والتي تكونت من المراحل الآتية : التحليل، البناء، التجريب، وتناولت الاختبارات ثلاثة مجالات رئيسية في الرياضيات (الهندسة، الحساب، القياس) حيث تكون اختبار الصف الرابع من (34) مفردة واختبار الصف الخامس من (37) مفردة واختبار الصف السادس من (32) مفردة.

نتائج الدراسة: بينت النتائج تمتع الاختبارات المحكية المرجع الثلاثة بدرجة جيدة من الصدق والثبات، وقدرتها على تصنيف الطلاب إلى متمكنين وغير متمكنين. كما أكدت النتائج أن هناك تدنياً واضحاً في درجة إتقان المهارات الأساسية في مادة الرياضيات لدى طلاب الصفوف العليا من المرحلة الابتدائية.

2- دراسة علام (2001):

عنوان الدراسة: اختبار تشخيصي مرجعي المحك في أساسيات القياس والتقويم التربوي والنفسي.

أهداف الدراسة: بناء اختبار تشخيصي مرجعي المحك لقياس تمكن طلاب كليات التربية والمعلمين في مصر الذين يتلقون تدريبا في مجال القياس والتقويم في أثناء الخدمة من أساسيات القياس والتقويم التربوي والنفسي

عينة الدراسة: عينة من طلاب كليات التربية والمعلمين الذين يتلقون تدريبا في مجال القياس والتقويم.

أداة الدراسة: اختبار تشخيصي مرجعي المحك من إعداد الباحث مؤلف من (60) مفردة من نوع الاختيار من متعدد بأربعة بدائل، ويقيس (57)، هدفا سلوكيا يتعلق بموضوعات محددة في مجال القياس والتقويم التربوي والنفسي، وقد اتبع الباحث الخطوات الآتية في بناء الاختبار:

1- تحديد محتوى الاختبار والأهداف السلوكية المتعلقة به.

2- بناء مفردات الاختبار والتحقق من صدقها: من خلال عرضها على مجموعة من المحكمين.

3- التحقق من ثبات الاختبار: باستخدام طريقة كودر ريتشارسون 21إذ بلغت قيمة معامل كودر(74.0) كماحسب الثبات وفق طريقة سبكوفياك إذ بلغت قيمة معامل الاتفاق لسبكوفياك (0.68).

4- تحديد الحد الأدنى لمستوى الإنجاز في الاختبار: باستخدام طريقة ندلسكاي، وطريقة إيبل، وطريقة المجموعات المتناقضة والمجموعات الحدية، وقد تراوحت درجة القطع وفق هذه الطرائق بين.(41-44).

نتائج الدراسة :بينت النتائج تمتع الاختبار بخصائص سيكومترية مقبولة، كما أن معامل الثبات المحسوب وفق معامل الاتفاق لسبكوفياك تعطي مؤشراً مقبولا لمدى فاعلية الاختبار في اتساق تصنيف الطلاب إلى متمكنين غير متمكنين (علام,2001 ،475).

3– دراسة عبدالله (2000).

عنوان الدراسة: القياس المحكي المرجع لأحد مقررات علم النفس التعليمي "دراسة ميدانية"

هدف الدراسة: بناء صور اختبارية محكية المرجع متعادلة القياس في أحد مقررات علم النفس التعليمي تقيس مدى تحقق الأهداف المعرفية المتعلقة بهذا المقرر لدى أفراد عيبة الدراسة.

عينة الدراسة: تمثلت عينة الدراسة في ثلاث شعب سحبت عشوائياً من شعب مجتمع الدراسة البالغ عددها خمس شعب تضم طالبات الفرقة الرابعة من الأقسام العلمية التربوية بكلية البنات جامعة عين شمس في مصر وبلغ حجم العينة (194) طالب

أداة الدراسة: الصورة الأولية لبنك الأسئلة والتي تقيس الأهداف الإجرائية لمادة علم النفس التعليمي المقررة على الفرقة الرابعة من الأقسام العلمية التربوية بكلية البنات جامعة عين شمس وقد نهجت الباحثة الخطوات الآتية لبناء أداة القياس (الصورة الأولية لبنك الأسئلة).

- تحديد المجال السلوكي الذي يقيسه بنك الأسئلة.
- بناء مفردات بنك الأسئلة التي تتضمن (إعداد جدول المواصفات وصياغة المفردات التي بلغ عددها (120) مفردة من نوع الاختيار من متعدد وصياغة التعليمات.
- الصدق الوصفي لأداة القياس وذلك من خلال صدق المحكمين وصدق البناء وصدق الملائمة كما يوفره نموذج راش وصدق المقارنة الطرفية.
- تم التحقق من ثبات الأداة بواسطة برنامج ميكروسكيل المستخدم في تدريج الاختبارات.

- **نتائج الدراسة:**

1- حذفت ست مفردات من مفردات البنك والبالغ عددها (120)، وذلك لعدم ملاءمتها لنموذج راش.

2- تم تعيين الدرجة الفاصلة المقابلة للتقديرات الجامعية المعتمدة، والتي تحدد محكات الإتقان على متصل القدرة وهذه الدرجات موحدة لجميع الصور الاختبارية، وتم على أساسها تحديد مستويات التحصيل لدى عينة الأفراد ومقارنة مستوى التحصيل لكل عينة من عينات الأفراد بكل محك من هذه المحكات.

3- تم تشخيص نقاط القوة والضعف في تحقيق أهداف القياس.

4- بينت الدراسة تحقق مستويات عليا من الأهداف في الفهم والتطبيق وعدم اقتصارها على التذكر.

4- دراسة الزهراني (2009):

عنوان الدراسة: بناء اختبار محكي المرجع لقياس الكفايات المهنية لدى عينة من معلمي المرحلة الابتدائية بمدينة الرياض.

هدف الدراسة: بناء اختبار محكي المرجع لقياس الكفايات المهنية لدى عينة من معلمي المرحلة الابتدائية بمدينة الرياض.

عينة الدراسة: تكونت عينة الدراسة من 38، معلماً تم اختيارهم بالطريقة العشوائية العنقودية من معلمي المرحلة الابتدائية بالرياض.

أداة الدراسة: قام الباحث ببناء اختار محكي المرجع مروراً بعدة خطوات هي:

1- مرحلة التحليل: وتم فيها تحديد المحتوى حيث تم تقسيم الكفايات المهنية إلى أربع كفايات رئيسية.

2- مرحلة البناء: وتم فيها إعداد مواصفات الأسئلة وذلك بصياغة الأهداف السلوكية ثم بناء مفردات لقياس تلك الأهداف، وذلك بالاعتماد على الخبراء في كل خطوة.

مرحلـة التجريـب: تـم فيهـا التحقـق مـن صـدق المقيـاس باسـتخدام صـدق المحتـوى بـالعرض علـى الخبـراء، وتـم التحقـق مـن الثبـات باسـتخدام معامـل لفنجسـتون، وتـم تحديـد درجة القطع لكل كفاية باستخدام طريقة أنجوف.

نتائج الدراسة :بعد المرور بتلك الخطوات تمثلت نتيجة الدراسة في إخراج الاختبار بصـورته النهائيـة والمكـون مـن سـتة اختبـارات فرعيـة تقيس أربـع كفايـات رئيسـة يفتـرض توفرها لـدى معلمـي المرحلـة الابتدائيـة، وبينـت الدراسـة الاختبـار يتمتـع الابتدائيـة، وبينـت الدراسة أن الاختبار يتمتع بدرجة مقبولة من الصدق والثبات.

- **الدراسات الأجنبيةك**

1- **دراسة كيث وقيرت (keith&Giert,2000)**

عنوان الدراسة: إجراءات تجميع اختبار مؤتمت لاختبار محكي المرجع

Automated Test Assembly procedures for Criterion Referenced Testing using optimization Heuristics

هدف الدراسة: بناء أشكال متوازية لاختبار محكي المرجع، والتحري عن كفاءة إجراء الاختبار المتسلسل الكترونياً وتأثيره.

عينة الدراسة: عينة عشوائية مكونة من (600) طالباً في لويزيانا.

أداة الدراسـة: بنـاء خمسـة اختبـارات محكيـة تحصـيلية لمقرر الرياضـيات فـي الصـف التاسع للولايات المتحدة الكترونياً، للعام الدراسي (1996/1995) وكل اختبار يتألف من (50) فقرة من نوع الاختيار من متعدد.

نتـائج الدراسـة: توصـلت إلـى أن اسـتخدام الاختبـارات المحكيـة المرجـع يحسـن مـن عمليـة القيـاس والتقويم، كمـا أن تخفيض طـول الاختبـار يزيـد مـن دقـة الاختبـار ويخفـض الكلفة والوقت في بناء الاختبارات، كما أنه يمكن استخدام الاختبارات المحكية المتوازية في تطوير بنوك الأسئلة.

2- دراسة دال(Dale,1995):

عنوان الدراسة: وضع اختبار لفصل دراسي للمعلمين الذين يكرهون وضع الاختبارات (بناء وتقويم الاختبار محكي المرجع)

Classroom Testing for Teachers who Hate Testing, Criterion-Referenced Test construction and Evaaluation

أهداف الدراسة: تقديم اختبارات محكية المرجع للمعلمين الذين يكرهون بشكل عام فكرة الاختبار أو الذين يتحاشون لسبب أو لآخر فضية الاختبار

عينة الدراسة: تكونت من 50 فرداً من طلاب وطالبات الفرقة الثانية بجامعة يابانية خاصة ضمن قسم منشأ حديثاً للدراسات الأمريكية الأوربية.

أداة الدراسة: اختبار TOEFL مكون من 50 مفردة ويتضمن ثلاث أنواع من المفردات وهي: اختيار من متعدد، صواب وخطأ، الإصغاء إلى فقرات مسجلة على أشرطة كاسيت

نتائج الدراسة:

- التمييز بين الاختبارات جماعية المرجع ومحكية المرجع، على أن الاختبارات جماعية المرجع قليلة الفائدة للمعلمين في تشخيص مواطن القوة والضعف لدى الطلاب وتقدير تقدمهم
- تحليل مفردات الاختبار يساعد المعلمين على تقدير ومراجعة وتحسين اختبارات صفوفهم.

ثالث عشر: نتائج الدراسة ومناقشتها:

السؤال الأول: ما مؤشرات صدق الاختبار المحكي المرجع في مادة الرياضيات؟

1- الصدق الداخلي:

استخدمت الباحثة لحساب الصدق الداخلي لاستبانة واقع تنفيذ الأنشطة البيئية إجابات 90 طالباً وطالبة شكلوا العينة الاستطلاعية للبحث، وذلك من خلال حساب

معاملات الارتباط بين كل فقرة من فقرات الاختبار والاختبار ككل، فإذا كان معامل الارتباط قوي ومعنوي إحصائياً دل ذلك على الصدق الداخلي للاستبانة.

الفقرة	معامل الارتباط مع الاختبار الكلي	مستوى الدلالة	القرار الاحصائي
.1	0.27	0.04	دالة
.2	0.33	0.01	دالة
.3	0.29	0.03	دالة
.4	0.43	0.001	دالة
.5	0.31	0.02	دالة
.6	0.4	0.002	دالة
.7	0.37	0.004	دالة
.8	0.39	0.003	دالة
.9	0.314	0.016	دالة
.10	0.30	0.021	دالة
.11	0.39	0.0036	دالة
.12	0.42	0.0036	دالة
.13	0.42	0.0033	دالة
.14	0.36	0.0032	دالة
.15	0.33	0.0042	دالة
.16	0.33	0.0034	دالة
.17	0.41	0.0039	دالة
.18	0.32	0.0041	دالة
.19	0.43	0.0037	دالة
.20	0.34	0.008	دالة
.21	0.34	0.0036	دالة
.22	0.36	0.0036	دالة
.23	0.3	0.0033	دالة
.24	0.29	0.0032	دالة
.25	0.37	0.0042	دالة
.26	0.33	0.0034	دالة
.27	0.35	0.0039	دالة
.28	0.43	0.001	دالة
.29	0.41	0.001	دالة
.30	0.29	0.013	دالة
.31	0.31	0.024	دالة
.32	0.34	0.005	دالة
.33	0.42	0.001	دالة

.34	0.33	0.009	دالة
.35	0.34	0.036	دالة
.36	0.36	0.042	دالة
.37	0.3	0.039	دالة
.38	0.29	0.037	دالة
.39	0.37	0.035	دالة
.40	0.33	0.034	دالة
.41	0.35	0.029	دالة
.42	0.3	0.0341	دالة
.43	0.33	0.033	دالة
.44	0.34	0.03	دالة
.45	0.31	0.034	دالة
.46	0.29	0.04	دالة
.47	0.36	0.003	دالة
.48	0.31	0.024	دالة
.49	0.35	0.037	دالة
.50	0.33	0.023	دالة
.51	0.36	0.007	دالة
.52	0.33	0.009	دالة
.53	0.38	0.0031	دالة
.54	0.33	0.014	دالة
.55	0.35	0.034	دالة
.56	0.3	0.04	دالة
.57	0.35	0.004	دالة
.58	0.34	0.0039	دالة
.59	0.36	0.003	دالة

تم التحقق من صدق الاختبار التمييزي والتلازمي ، وفيما يلي عرض لذلك:

3- الصدق التمييزي: جرى ترتيب أفراد العينة وفق الدرجة الكلية للاختبار ومن ثمّ اختيار المجموعة العليا تحوي /27%/ من الطلاب أي/ 145 /طالباً وكذلك الأمر بالنسبة للمجموعة الدنيا، تم بعد ذلك تطبيق اختبار t-test للمقارنة بين متوسطي درجات المجموعتين وكانت النتيجة.

الجدول (1): اختبار t-test للتحقق من الصدق التلازمي للاختبار المحكي

المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة t	درجة الحرية	مستوى الدلالة	القرار
العليا	145	44.68	8.67	32.666	288	0.000	دال إحصائياً عند 0.01
الدنيا	145	20.32	2.34				

يتضح من الجدول أعلاه وجود فروق ذات دلالة إحصائية بين متوسط درجات المجموعتين على الاختبار ما يدل على قدرة الاختبار على التمييز بين المستويات المختلفة للطلاب.

3- الصدق التلازمي:

تم حساب معامل ارتباط بيرسون بين درجات الطلبة على الاختبار ودرجاتهم في اختبار الامتحان النهائي لمادة الرياضيات للفصل الدراسي الثاني للعام الدراسي 2018/ 2019م، وقد بلغت قيمته (0.841) بمستوى دلالة (0.000)، وهو معامل ارتباط مرتفع ودال إحصائياً عند مستوى دلالة (0.01)، ما يشير إلى تمتع الاختبار بالصدق التلازمي.

السؤال الثاني: ما مؤشرات ثبات الاختبار المحكي المرجع في مادة الرياضيات ؟

تم التحقق من ثبات الاختبار وفق عدة طرائق هي:

1- الثبات باستخدام معامل كودر–ريتشارد سون–21:

تم استخدام معامل كودر ريتشارد سون (21) لحساب الثبات لأن البيانات من النوع الثنائي (0–1) وهي معامل ثبات قريب من معامل ألفاكرونباخ والذي أعطى قيمة /0.89/ وهي قيمة ثبات عالية. تبين أن الاختبار يتمتع ببنية داخلية ثابتة.

2- الثبات باستخدام التجزئة النصفية:

تم أيضا التحقق من ثبات التجزئة النصفية من خلال حساب معامل ارتباط سبيرمان-براون المعدل لمعامل ارتباط بيرسون بين جزئي الاختبار بعد تقسيم الاختبار إلى قسمين يحوي القسم الأول البنود الزوجية ويحوي الجزء الثاني البنود الفردية وكانت قيمة معامل ارتباط سبيرمان-براون /0.740/ وهو معامل ثبات جيد.

عدد البنود	معامل ثبات التجزئة النصفية
77	0.740

3- الثبات بالإعادة:

وهو معامل الارتباط بين الدرجات التي يحصل عليها مجموعة المفحوصين نفسها على التطبيق الأول والتطبيق الثاني للمقياس نفسه (عباس،1996، ص43).

وللتحقق من ثبات الإعادة للاختبار قامت الباحثة بإعادة التطبيق على العينة نفسها بعد مرور (خمسة عشر) يوماً وبلغ عدد الطلاب الذين تم تطبيق الاختبار عليهم في المرتين (432) طالباً وطالبةً، وقد تم حساب معامل الارتباط بين درجاتهم في المرتين وبلغ (0.88) وهو معامل ارتباط مرتفع ودال إحصائياً عند (0.01).

عدد الطلاب	معامل ارتباط بيرسون	القيمة الاحتمالية	القرار
580	0.88	0.00	دال

السؤال الثالث: ما مؤشرات القدرة التمييزية للاختبار المحكي المرجع في مادة الرياضيات ؟

جرى حساب معاملات التمييز للبنود بعد ترتيب درجاتهم من الأعلى للأدنى وأخذ المجموعتين العليا والدنيا بنسبة (27%) لكل مجموعة، ومن ثم حساب معاملات التمييز للمجموعة العليا، وللمجموعة الدنيا، والجدول الآتي يوضح معاملات التمييز لكل بند:

الجدول(2): معاملات تمييز بنود الاختبار المحكي المرجع

البند	معامل التمييز	البند	معامل التمييز	البند	معامل التمييز	البند	معامل التمييز
a1	560.	**a17**	0.46	**a33**	0.14	**a49**	0.56
a2	0.29	**a18**	0.26	**a34**	0.43	**a50**	0.32
a3	0.24	**a19**	0.48	**a35**	0.38	**a51**	0.38
a4	0.53	**a20**	0.64	**a36**	0.40	**a52**	0.33
a5	0.53	**a21**	0.59	**a37**	0.37	**a53**	0.55
a6	0.27	**a22**	0.39	**a38**	0.42	**a54**	0.33
a7	0.38	**a23**	0.26	**a39**	0.43	**a55**	0.39
a8	0.45	**a24**	0.49	**a40**	0.56	**a56**	0.49
a9	0.49	**a25**	0.26	**a41**	0.47	**a57**	0.34
a10	0.40	**a26**	0.55	**a42**	0.37	**a58**	0.52
a11	0.63	**a27**	0.31	**a43**	0.34	**a59**	0.37
a12	0.38	**a28**	0.39	**a44**	0.39		
a13	0.56	**a29**	0.31	**a45**	0.45		
a14	0.20	**a30**	0.36	**a46**	0.26		
a15	0.30	**a31**	0.45	**a47**	0.24		
a16	0.25	**a32**	0.23	**a48**	0.37		

يظهر من الجدول أعلاه أنّ معاملات التمييز للبنود قد كانت ضمن الحدود المقبولة وتراوحت بين (0.23 و 0.63) باستثناء البندين: (14، 33) إذ كان معامل التمييز لهما منخفض.

القدرة التمييزية:

تم حساب الاجابات الصحيحة للفقرات ل 27% من العينة ذو الاجابات الاعلى وكذلك ل27% من العينة أصحاب الإجابات الأدنى حيث تم حساب عدد الأفراد الذين أجابوا إجابة صحيحة على الفقرة في كل من المجموعتين ثم نقوم بحساب النسبة المئوية لدرجة التمييز وفق ما يلي:

القدرة التمييزية=100 ×(عدد الاجابات الصحيحة لأعلى 27%- عدد الاجابات الصحيحة لأدنى 27%) ÷ عدد الفئة الواحدة.

واعتمدت الباحثة على المعيار الآتي:

القيم السالبة للفقرات تدل على أن الفقرات غير مميزة

القيم الموجبة الأدنى من 40 % ذات تمييز ضعيف

القيم الموجبة بين(40-60) % ذات تمييز متوسط

القيم التي أكبر من 60% ذات تمييز جيد

الفقرة	عدد الاجابات الصحيحة للفئة الاعلى	عدد الاجابات الصحيحة للفئة الادنى	القدرة التمييزية %
.1	19	6	48
.2	18	6	44
.3	15	4	41
.4	16	4	44
.5	15	7	30
.6	19	5	52
.7	16	4	44
.8	17	4	48
.9	17	3	52
.10	17	2	56
.11	19	4	56
.12	19	5	52
.13	16	5	41
.14	19	5	52
.15	19	6	48
.16	16	2	52
.17	15	2	48
.18	16	2	52
.19	15	2	48
.20	18	1	63
.21	19	3	59
.22	18	5	48
.23	16	4	44
.24	18	4	52

44	3	15	.25
52	1	15	.26
59	3	19	.27
48	3	16	.28
56	3	18	.29
48	2	15	.30
41	6	17	.31
41	4	15	.32
48	6	19	.33
48	3	16	.34
41	8	19	.35
41	7	18	.36
48	3	16	.37
48	3	16	.38
44	5	17	.39
48	2	15	.40
63	1	18	.41
67	1	19	.42
48	2	15	.43
56	2	17	.44
59	2	18	.45
56	2	17	.46
30	10	18	.47
59	1	17	.48
52	3	17	.49
41	4	15	.50
44	3	15	.51
48	2	15	.52
52	1	15	.53
59	2	18	.54
59	3	19	.55
56	3	18	.56
56	4	19	.57
48	4	17	.58
56	1	16	.59

ان معظم الفقرات الباقية ذات قدرة تمييزية متوسطة وبالتالي الاختبار جاهز للتطبيق

السؤال الرابع– ما مؤشرات سهولة البنود للاختبار المحكي المرجع في مادة الرياضيات ؟

جرى حساب معاملات السهولة لكل بند من خلال حساب النسبة المئوية للطلاب الذين أجابوا إجابة صحيحة عن السؤال، والجدول الآتي يوضح ذلك:

الجدول (3): معاملات السهولة لبنود الاختبار المحكي المرجع

البند	معامل السهولة	البند	معامل السهولة	البند	معامل السهولة	البند	معامل السهولة
a1	0.70	a17	0.60	a33	0.38	a49	0.45
a2	0.75	a18	0.76	a34	0.76	a50	0.63
a3	0.65	a19	0.38	a35	0.64	a51	0.27
a4	0.68	a20	0.60	a36	0.51	a52	0.66
a5	0.68	a21	0.62	a37	0.42	a53	0.34
a6	0.77	a22	0.66	a38	0.69	a54	0.45
a7	0.39	a23	0.74	a39	0.38	a55	0.51
a8	0.75	a24	0.48	a40	0.43	a56	0.77
a9	0.75	a25	0.73	a41	0.31	a57	0.24
a10	0.69	a26	0.38	a42	0.78	a58	0.47
a11	0.60	a27	0.45	a43	0.41	a59	0.34
a12	0.64	a28	0.47	a44	0.76		
a13	0.46	a29	0.24	a45	0.35		
a14	0.39	a30	0.76	a46	0.84		
a15	0.75	a31	0.45	a47	0.34		
a16	0.79	a32	0.38	a48	0.69		

يفيد معامل السهولة في إيضاح مدى سهولة بند ما في الاختبار، وهو عبارة عن النسبة المئوية من الطلاب الذين أجابوا عن السؤال إجابة صحيحة، فكلما زاد عدد الإجابات الصحيحة للبند ارتفع معامل السهولة وقل معامل الصعوبة، وكلما انخفضت النسبة المئوية للإجابات عن الأسئلة أصبح البند سهل جداً وبالتالي يجب حذفه كما في البند (7) مثلاً حيث معامل سهولته 39% وهو خارج الحدود المقبولة أي بين (0.60– 0.80).

بالعودة للجدول السابق يتضح أنّ معاملات السهولة للبنود كانت ضمن الحدود المقبولة ماعدا البنود الآتية (7، 13، 14، 19، 24، 26، 27، 28، 29، 31، 32، 33، 36، 37، 39، 40، 41، 43، 45، 46، 47، 49، 51، 53، 54، 55، 57، 58، 59). كانت معاملات سهولتها ضعيفة تم حذفها.

وتوصلت الدراسة إلى بناء اختبار محكي المرجع مكون من (59) بصورته النهائية ويتمتع بخصائص سيكومترية جيدة ويمكن الاستفادة منه وتطبيقه من قبل المعلم و بشكل مباشر على الطلاب.

رابع عشر: مقترحات البحث:

- بناء اختبارات محكيه المرجع جاهزة للاستخدام من قبل المعلمين.
- بناء بنوك أسئلة في المقررات المختلفة وتقنينها وتدريب المعلمين على استخدامها.
- إقامة دورات تدريبية مستمرة على بناء الاختبارات باستخدام البرمجيات المناسبة.

المراجع:

المراجع العربية:

1- أبو سل، محمد (2002). قياس وتقويم تعلم الطلبة. الأردن، عمان: دار الفرقان

2- راشد، محمد (2005). مناهج الرياضيات وأساليب تدريسها للصفوف الرئيسية، الأردن، دار الجنادرية.

3- رضوان، رانية جميل (2011). أثر طرائق تقدير الدرجات في الخصائص السيكومترية لاختبار تحصيلي ذي الاختيار من متعدد في مادة الرياضيات للصف الرابع من التعليم الأساسي. رسالة ماجستير، قسم القياس والتقويم التربوي والنفسي، كلية التربية، كلية التربية، جامعة دمشق

4- زهراني، سامي صالح سرحان.(2001). بناء اختبار محكي المرجع لقياس الكفايات المهنية لدى عينة من معلمي المرحلة الابتدائية بمدينة الرياض، رسالة ماجستير، قسم علم النفس، جامعة الملك سعود، الرياض، المملكة العربية السعودية.

5- سنقر، صالحة. (2000). الطرائق الخاصة في التعليم الابتدائي، دمشق، منشورات جامعة دمشق.

6- عبابنة، عماد غصاب.(2009).الاختبارات محكية المرجع فلسفتها وأسس تطويرها، ط1، الأردن، دار الميسرة للنشر.

7- عبدالله، اعتدال غازي. (200). القياس محكي المرجع لأحد مقررات علم النفس التعليمي "دراسة ميدانية" . رسالة ماجستير، قسم علم النفس، كلية البنات للآداب والعلوم والتربية، جامعة عين شمس، القاهرة، مصر.

8- علام، صلاح الدين محمود.(2006). القياس والتقويم التربوي والنفسي، القاهرة، دار الفكر العربي.

9- علام، صلاح الدين محمود.(2001). الاختبارات التشخيصية مرجعية المحك، القاهرة، دار الفكر العربي.

10– علام، صلاح الدين محمود.(2000). القياس والتقويم التربوي والنفسي، القاهرة، دار الفكر العربي.
11– علام، صلاح الدين محمود.(2005). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي، القاهرة، دار الفكر العربي.
12– العنزي، حسين بن عزيز حسين(2004). بناء اختبار تحصيلي لقياس درجة اتقان المهارات الأساسية في مادة الرياضيات لدى تلاميذ الصفوف العليا في المرحلة الابتدائية. رسالة ماجستير، كلية التربية، جامعة الملك سعود، الرياض، المملكة العربية السعودية
13– عودة، أحمد.(2010). القياس والتقويم في العملية التدريسية، ط1، الأردن، دار الأمل للنشر.
14– عودة، أحمد.(1993). القياس والتقويم في العملية التدريسية، ط2، الأردن، دار الأمل للنشر.
15– فرج، صفوت.(2007). القياس النفسي، ط6، القاهرة، مكتبة الأنجلو المصرية.
16– كاظم، أمينة.(1988). دراسة نظرية نقدية حول المقياس (نموذج راش)، مؤسسة الكويت للتقدم العلمي، سلسلة الكتاب المتخصصة.
17– مراد، صلاح، سليمان، أمين.(2002). الاختبارات والمقاييس في العلوم النفسية والتربوية خطوات إعدادها وخصائصها، ط1، القاهرة/، دار الكتاب الحديث.
18– ميخائيل، امطانيوس.(2012). القياس والتقويم في التربية الحديثة، منشورات جامعة مشق، كلية التربية.
19– ميخائيل ، امطانيوس.(2010).القياس والتقويم في التربية الحديثة، منشورات جامعة دمشق، كلية التربية.
20– ميخائيل، امطانيوس. (2003). القياس والتقويم في التربية الحديثة، جامعة دمشق، كلية التربية.

المراجع الأجنبية:

1- Anastasi, anne.(1982).Psychological Testing (fifth Edition). Mlacmillan Publishing co.Inc. Newyork.
2- Brian.D. (2009). Criterion Referenced test is an a mastery test.By: www.mountain measurement.com
3- Dale,G.T.(1995): Classrom Testirg for teachers who Hate testing criterion Referenced Test, construction and Evaluation, Reports-Research, Techincal(143) us department of Education office of Educational Resarch and Imporvement (OERI) Washington, p, 1-20.
4- Keith,A & Giert. J. (2000). Automated b Test Assembly Procedures for Criterion Referenced Testing Using Optimization Heuristics .Paper Presented at the Annual Meeting of the AERA. NEW Orleans Louisiana, USA.
5- Hambleton,Ronald(2009) Criterion- Refrrenced Test. By: http://www.education.com
6- Thorndikc, R, M. ct. al.(1991). measurement and evaluation in psychology and education. New York, Macmillan publishing company.
7- Popham, W, James (2014). Criterion-Referenced Measurement: Half a Century Wasted? Educational Leadership, 7(6), 62-66.