Journal of Educational and Psychological Sciences
Volume (4), Issue (40) : 30 Oct 2020
P: 179 - 194

AJSRP
ISSN: 2522-3399

مجلة العلوم التربوية والنفسية
المجلد (4)، العدد (40) : 30 أكتوبر 2020 م
ص: 179 - 194

Constructing an Achievement Reference Test to measure language competencies – A descriptive survey study on secondary school students in Riyadh City –

Ibrahim Z. ALAssiri
Riyadh Education Directorate || Ministry of Education || KSA
Mohammed R. ALZahrani
Faculty of Education || Umm Al-Qura University Makkah || KSA

Abstract: The present study aims to build a criterion- reference test to measure linguistic competencies of secondary school students in Riyadh, and to verify its psychometric characteristics. The descriptive survey method was used as it is the appropriate method to achieve the objectives of the study. Further, the sample of the study consists of (50769) students, representing all secondary students in Riyadh between 1439/1440 AH in accordance with the data obtained by the researcher. Also, the population of the current study consists of (983) secondary students in Riyadh as a sample. The researcher built the test in four stages: (analysis stage- construction phase- experimentation stage- output stage). After passing these steps to build the test, the study produced a criterion-reference test to measure the linguistic competencies for the secondary school students which consist of (40) questions measuring two linguistic competencies (syntactic and dictation competency).

The study found that this test has good validity and it was found that the test has high reliability according to Livingstone coefficient to the degree of (85%). Based on these results, the study recommends that the application of the test built in the current study is appropriate to measure students' linguistic competencies. This study also recommends that the measurement and evaluation experts should find new ways to build tests based on criterion-reference in all subjects of education.

Keywords: Achievement test – criterion-reference test- linguistic competencies- secondary school – Riyadh.

بناء اختبار تحصيلي محكي المرجع لقياس الكفايات اللغوية – دراسة وصفية مسحية على طلاب المرحلة الثانوية بمدينة الرياض –

إبراهيم زايد العسيري
إدارة تعليم الرياض || وزارة التعليم || المملكة العربية السعودية
محمد رزق الله الزهراني
كلية التربية || جامعة أم القرى || مكة المكرمة || المملكة العربية السعودية

الملخص: هدفت الدراسة الحالية إلى بناء اختبار محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض، والتحقق من خصائصه السيكومترية. وتم استخدام المنهج الوصفي المسحي؛ كونه المناسب لتحقيق أهداف الدراسة، وتكون مجتمع الدراسة الحالية من جميع طلاب المرحلة الثانوية بمدينة الرياض للعام الدراسي 1440/1439هـ والبالغ عددهم (50769) حسب الإحصاءات التي تحصل عليها الباحث، كما تكونت عينة الدراسة من (983) طالباً من المرحلة الثانوية في مدينة الرياض، وقد قام

الباحث ببناء الاختبار مروراً بأربع مراحل وهي (مرحلة التحليل – مرحلة البناء – مرحلة التجريب – مرحلة الإخراج) وبعد المرور بهذه الخطوات لبناء الاختبار أنتجت الدراسة اختباراً محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية يتكون من (40) مفردة اختبارية تقيس كفايتين لغويتين وهي (الكفاية النحوية والكفاية الإملائية).

وتوصلت الدراسة إلى أنَّ هذا الاختبار يتمتع بصدق جيد كما تبين أنَّ الاختبار يتمتع بثبات جيد وفقاً لمعامل ليفنجستون وصل إلى درجة (85%)، وبناءً على هذه النتائج أوصت الدراسة بتطبيق الاختبار الذي تم بناؤه في الدراسة الحالية على الطلبة؛ لتشخيص مدى تمكنهم في الكفايات اللغوية، وتوصية خبراء القياس والتقويم بإيجاد طرق جديدة لبناء الاختبارات محكية المرجع في كافة مواضيع التربية.

الكلمات المفتاحية: اختبار تحصيلي، محكي المرجع، الكفايات اللغوية، المرحلة الثانوية، الرياض.

المقدمة

تعتمد عملية القياس التربوي والنفسي على أدوات ومقاييس تقدم بيانات كمية تتيح للتربويين فهم الظاهرة التربوية بشكل دقيق، وتعد الاختبارات التحصيلية جزءًا من هذه الأدوات فمن خلالها يمكن الحكم على مدى تحقق الأهداف التعليمية كما تفيد في التعرف على مستوى الطالب وتحصيله الدراسي.

ومن خلال متابعة التطورات في بناء المقاييس التربوية والنفسية نلاحظ اتجاهًا نحو استخدام قياس معياري المرجع وبالمقابل هناك اتجاه آخر نحو استخدام القياس محكي المرجع، وتجدر الإشارة إلى أنَّ كثيراً من الدراسات تناولت أوجه الشبه والاختلاف بين القياس المحكي المرجع والمعياري المرجع، فالقياس معياري المرجع يعتمد على تفسير الدرجات في ضوء معايير محددة، أي تنسب درجات الطالب في الاختبار إلى جداول معيارية، ومن ثم تحديد مستوى أداء الطالب النسبي بناءً على موقع درجته مقارنة بمتوسط تحصيل المجموعة المعيارية في الاختبار، ويركز هذا الاتجاه على الفروق الفردية بين الطلبة بحيث يوضح مدى الفارق بين مستوى طالب وآخر أي أنه يقيس مقدرة طالب معين مقارنة بمقدرة طلبة آخرين. وتهتم هذه الاختبارات المعيارية المرجع بالفروق الفردية بين الطلبة دون النظر إلى مدى إتقان الطالب للمهارات والمعلومات المراد قياسها، وبالتالي فهي لا تساعد على تقويم وتشخيص كل كفاية. كما أنها لا تزود المعلم بمعلومات دقيقة تساعد في اتخاذ القرارات التربوية المناسبة، وهذه الانتقادات لعبت دوراً في ظهور تصورات حديثة في منهجية الاختبارات. والمقاييس التربوية أدت إلى ظهور اتجاه آخر أطلق عليه القياس محكي المرجع. (علام، 2001).

وعلى الرغم من أهمية الاختبارات محكية المرجع في تشخيص الأداء وقياس الكفايات إلا أنه يلاحظ قلة هذه الاختبارات بل ندرتها في الوطن العربي؛ إذ يوكل أمر إعدادها في الغالب إلى مدرسين لا تتوافر لديهم الإمكانيات والقدرات اللازمة لإعدادها. ومع ذلك ظهرت محاولات عربية لبنائها، إلا أنَّ هذه المحاولات ما زالت متواضعة، أو أنها لم تخرج من المنطقة التي أجريت فيها، فغالباً ما يتم بناء اختبارات تشخيصية محكية المرجع لمهمة محدودة، وليس لتطوير استراتيجية لبناء هذا النوع من الاختبارات، الأمر الذي يجعل مدخل الاختبارات محكية المرجع مجالاً خصباً للدراسات والبحوث المستقبلية (الحموري والكحلوت، 2009).

ويرجع البعض ندرة توفر الاختبارات محكية المرجع في الوطن العربي إلى ندرة الخبرات العربية في مجال تطوير هذا النوع من الاختبارات، من حيث معرفة كيفية بنائها، وتفسيرها. كما أنها تتطلب كلفة مادية وجهدين كبيرين (علام، 2000).

نظراً لما للكفايات اللغوية من أهمية بالغة في إكساب الطالب القدرة على التواصل اللغوي السليم، بحيث يصبح مستمعاً واعياً، ومتحدثًا بليغاً، وقارئاً وكاتباً مجيداً، فإذا تمكن متعلم اللغة العربية من تحقيق ما سبق، فإنه سيصبح متمكناً من مهارات اللغة الأربع (الاستماع والتحدث، والقراءة، والكتابة)، الأمر الذي أكد عليه كل من

(مدكور، 1991؛ والجليدي، 2011) بأنه ينبغي أنْ تكون كل محاولة لتعليم اللغة العربية هي السعي لتحقيق هذا الهدف وصولاً إلى التمكن من مهاراتها.

كما تُعد أساليب التقويم في تدريس مقرر الكفايات اللغوية أمرًا مهماً لقياس أداء المتعلمين إذ يمثل التقويم مرحلة مهمة في الربط بين الممارسات التعليمية وتخطيطها لذا سعى الباحثان من خلال دراسته هذه للتوصل إلى أداة تقويمية تسهم في التأكد من مدى توافر هذه الكفايات اللغوية لدى طلاب المرحلة الثانوية.

وبناءً على ما سبق فقد قام الباحثان ببناء اختبار تحصيلي محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض، وتبنى في ذلك الخطوات العلمية المتعارف عليها عند متخصصي القياس النفسي والتربوي في بناء هذا الاختبار التشخيصي، ثم القيام بتحليل نتائجه وتفسيرها بهدف التوصل إلى خصائصه السيكومترية، وذلك في ضوء فلسفة القياس محكي المرجع.

مشكلة البحث

أكدت العديد من الدراسات وجود ضعف في التحصيل اللغوي لطلاب المرحلة الثانوية بالوطن العربي كدراسة أحمد (2009)، وعوض (2010). كما أكدت العديد من الدراسات المحلية وجود ضعف واضح في خريجي التعليم الثانوي، يتركز ضعفهم في إتقان اللغة العربية، وعدم قدرتهم على الاستنتاج والاستدلال نظراً لتعودهم على الحفظ والاستماع، ويتجلى ذلك في القدرات المقاسة في الجزء اللفظي في اختبار القدرات العامة للطلاب.(الحربي، 2013).

وبعد إطلاع الباحثين على طرق تقويم طلاب المرحلة الثانوية في مقررات الكفايات اللغوية تبين عدم توفر معايير ثابتة تستند على أسلوب علمي موضوعي دقيق في تقويم أدائهم، ويمكن أنْ يكون ذلك سبباً في ضعف مستوى الطلاب، ومن خلال ملاحظة الباحث تبين له عدم وجود أداة تتسم بخصائص سيكومترية جيدة لقياس تحصيلهم يمكن من خلالها التحقق من مدى إتقانهم للمقررات الدراسية، وخاصة الكفايات اللغوية التي لا بد من توافرها في الطالب عند التخرج من المرحلة الثانوية.

وحيث إنَّ العملية التعليمية في التعليم الثانوي تحتاج إلى استخدام الاختبارات المحكية التي تبين مدى التقدم والتحسن لدى الطلاب، كل ما سبق دعا الباحث إلى إجراء هذه الدراسة وهي: بناء اختبار تحصيلي محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض.

أسئلة البحث

ويمكن تحديد مشكلة الدراسة في التساؤلات التالية:

1- هل يتوافر للاختبار محكي المرجع- الذي تم بناؤه- والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض مؤشر مقبول من الصدق؟
2- هل يتوافر للاختبار محكي المرجع- الذي تم بناؤه- والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض درجة مقبولة من الثبات؟
3- هل يتوافر للاختبار محكي المرجع- الذي تم بناؤه- والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض درجة مقبولة من التمييز؟
4- هل يتوافر للاختبار محكي المرجع- الذي تم بناؤه- والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض درجة مقبولة من الصعوبة ؟

أهداف البحث

تستهدف الدراسة الحالية:

1- بناء اختبار تحصيلي محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية في مدينة الرياض، والتحقق من خصائصه السيكومترية.
2- التعرف على مدى توافر الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض.

أهمية البحث

تكمن أهمية الدراسة الحالية في النقاط التالية:

1- توفير اختبار تحصيلي محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية
2- تزويد المسئولين في وزارة التعليم بتصور موضوعي متكامل عن مستوى الكفايات اللغوية لدى طلاب المرحلة الثانوية.
3- مساعدة المعلمين على معرفة الكفايات اللغوية الواجب توافرها لدى طلاب المرحلة الثانوية، كما يمكنهم استخدام هذا الاختبار في التقويم الختامي للطلاب، وبالتالي العمل على تطوير أدائهم من خلال معرفتهم لجوانب القصور لدى الطلاب.

2- الإطار النظري والدراسات السابقة

خطوات بناء الاختبارات محكية المرجع:

تختلف خطوات بناء الاختبارات محكية المرجع عن خطوات بناء الاختبارات معيارية المرجع، وذلك بسبب اختلاف هدف كل نوع منهما، والقرارات التي تتمخض عن نتائجهما.

فالاختبارات معيارية المرجع كما تبين سابقاً تهدف إلى مقارنة أداء الفرد بأداء المجموعة المعيارية، من خلال مقارنة درجته بمتوسط أداء المجموعة لتحديد موقعه بالنسبة لهذه المجموعة. والدرجة في هذا النوع تتسم بالعمومية والشمول للمجال الذي تقيسه، فهي لا ترمز على جزئيات المجال المقاس(علام، 2001، أ).

أما الاختبارات محكية المرجع فتهدف إلى مقارنة أداء الفرد بمحك أداء مسبق، لمعرفة ما حققه الفرد وما فشل في تحقيقه في البرنامج التعليمي، وتقدم وصفاً تفصيلياً لحالة الفرد بالنسبة للمحك أو المجال السلوكي الذي يقيسه. فهي تساهم في تشخيص جوانب القوة والضعف عند المتعلم وفي البرامج التعليمية.

وعند بناء الاختبارات محكية المرجع هناك خطوات هامة ينبغي المرور بها إذ إنَّ بناء هذا النوع من الاختبارات يتطلب الرجوع إلى المصادر والمراجع المتخصصة بالاختبارات محكية المرجع بالإضافة إلى الدراسات السابقة والاستفادة من خطواتها.

وعند مسح التراث الأدبي لهذا النوع من الاختبارات نجد أنَّ هناك مجالين يتم فيهما بناء اختبارات محكية وهما (المجال المعرفي والمجال الوجداني) وداخل كل مجال من المجالين توجد خطوات مختلفة يجب إتباعها عند بناء الاختبارات محكية المرجع. أوردها (علام، 2001؛ أ؛ الطريري، 2009) على النحو التالي:

الخطوة الأولى: تحديد النطاق السلوكي المراد قياسه

ويقصد بالنطاق السلوكي " مجموعة المعارف والمهارات المحددة تحديداً دقيقاً والمتعلقة بمجال تعليمي معين مما يمكننا من معرفة ما يستطيع الفرد أداءه وما لا يستطيع أداءه من خلال – الاختبار". وتختلف النطاقات السلوكية في درجة تعقدها ونوعية الأداء الذي تتضمنه (علام، 2000).

الخطوة الثانية: تحليل النطاق السلوكي على الكفايات الفرعية (علام، 2001، أ).

وفيها يتم تحليل النطاق السلوكي إلى الكفايات الفرعية؛لأنَّ الكفايات والمهارات الرئيسة تعد بمثابة نواتج تعلم مركبة وتتضمن معارف وعمليات عقلية ونفسية وحركية، لذلك فهي تتطلب تحليلاً إجرائياً أو بنائياً وترتيب مكوناتها ترتيباً منطقياً يكشف عن العلاقة القائمة بينهما.

الخطوة الثالثة: صياغة الأهداف السلوكية المتعلقة بنواتج تحليل النطاق السلوكي

ويتم تحديد الأهداف السلوكية لكل كفاية عن طريق استمارة تحكيم يعدها معد الاختبار ثم يتم عرضها على المحكمين في مجال الكفايات المقاسة، ثم يتم حساب معامل الاتفاق بين المحكمين بالمعادلة التالية:

عدد المحكمين × عدد الأهداف المتفق على أنها ملائمة

معامل الاتفاق = ——————————————— ... (1)

ن1 + ن2 +..... ن

حيث إن:

ن1 = عدد الأهداف التي صنفها المحكم الأول على أنها ملائمة.

ن2= عدد الأهداف التي صنفها المحكم الثاني على أنها ملائمة..... وهكذا. (الصبحي، 2001)

الخطوة الرابعة: بناء الاختبار

يتم انتقاء أنسب أنواع المفردات التي تقيس الأهداف السلوكية المحددة قياساً مباشراً، ومهارة كبيرة في فنيات بناء هذه المفردات، كما يتطلب التمكن من محتوى البرنامج التعليمي أو التدريبي، وفهم خصائص المتعلمين وذلك؛ لأنَّ هذه المفردات تستخدم في التمييز بين الذين استطاعوا تحقيق الأهداف والذين واجهتهم صعوبات في تحقيقها (مجيد، 2007).

الخطوة الخامسة: جميع مفردات الاختبار

بعد الانتهاء من كتابة المفردات وتدقيقها يبدأ الباحث بتنسيقها في ضوء الأهداف التي تقيسها لتكوين الاختبار، ويتطلب هذا مراعاة التسلسل المنطقي للمفردات وتجميع المفردات في ضوء مكونات المهارة أو الكفاية التي تم تحليل مستوياتها المعرفية، مع توضيح كيفية الإجابة ومكانها وتحديد زمن الاختبار (علام، 2001، أ).

الخطوة السادسة: تجريب الاختبار ميدانيًا وتحليل مفرداته

ذكر علام (2001، ب) أنَّ التجريب يساعد في التعرف على المشكلات الموجودة في الأسئلة سواء كان ذلك من حيث الصياغة أو الأسلوب أو الوضوح أو مستويات الصعوبة والتمييز، كما يساعد في تقدير الوقت اللازم للاختبار وعدد الأسئلة التي تلزم كتابتها في ضوء الوقت المتاح.

الخطوة السابعة: التحقق من صدق دلالات الاختبار وثباته

يقصد بالصدق مدى ملائمة الدرجات المستمدة من الاختبار للاستخدامات المعينة المناسبة للغرض الذي بني من أجله الاختبار، ويزداد ثبات الاختبار بزيادة هذا الاتساق، فالاختبار محكي المرجع يجب أنْ يتمتع بالثبات شأنه في ذلك شأن الاختبار معياري المرجع، غير أنَّ الثبات هنا يعني اتساق قياس نطاق سلوكي معين، وإذا كان الاختبار يقيس أكثر من مهارة واحدة فإنه يجب تقدير ثبات كل اختبار فرعي يقيس إحدى هذه المهارات (علام، 2001، أ)،

ثانياً- الدراسات السابقة:

وقد تناولت العديد من الدراسات في البحوث العلمية بناء الاختبارات محكية المرجع، وفي المكتبة العربية حصل الباحثان على العديد منها، وبما أنَّ هدف الدراسة الحالية هو بناء اختبار تحصيلي محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض، فإنه سيتم عرض الدراسات السابقة على النحو التالي:

أولاً- دراسات تناولت بناء الاختبارات محكية المرجع.

- **دراسة التميمي (1999)** والتي تمثل هدفها في تحديد أثر كل من نوع المحكم وطول الاختبار على درجة القطع لاختبار محكي المرجع يقيس الكفايات الرياضية في العمليات الحسابية على الأعداد بالصف السادس الابتدائي بمدينة جدة، وأجريت الدراسة على عينة من طلاب الصف السادس الابتدائي وبلغ عددهم (659) طالباً، وعينة من معلمي الرياضيات بلغ عددهم (20) معلماً لمادة الرياضيات، وعينة من مشرفي مادة الرياضيات بلغ عددهم (10) أعضاء، وكان التساؤل الرئيس للدراسة: هل تختلف درجة القطع في الاختبار محكي المرجع عند تطبيق طرق تحديد درجة القطع الأربعة على اختبار موحد وعينة دراسية واحدة باختلاف نوع المحكم وطول الاختبار؟ وأظهرت النتائج عدم وجود فروق دالة إحصائياً بين درجات قطع الاختبار الناتج باختلاف نوع المحكم وفق طريقة أنجوف.
- **ودراسة الشريفين (2006)** تمثل هدفها في تقدير الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق نظرية القياس الحديثة، وتكونت العينة من (222) طالباً وطالبةً من طلاب جامعة آل البيت موزعين على (4) شعب، ومسجلين في مساق القياس، والتقويم التربوي، طبق عليهم لاختبار تحصيلي في القياس والتقويم المؤلف من (50) فقرة من نوع الاختيار من متعدد ذي الأربعة بدائل، كما تم استخدام أساليب إحصائية معامل التمييز ومعامل الصعوبة، وأظهرت النتائج مطابقة الاستجابات عن (37) فقرة من فقرات الاختبار لافتراضات نموذج راش، وحذف (13) لم يطابق النموذج، كما بلغ معامل الثبات للأفراد (0.90)، ومعامل الثبات للاختبار (0.95)، وتم التأكد من تمتع الاختبار بمظاهر متعددة من الصدق ومنها صدق المحتوى، وأنَّ الاختبار يقدم أكبر كمية من المعلومات عند مستويات القدرة المتوسطة، ويقدم أقل كمية من المعلومات عند مستويات القدرة العالية والمتدنية.
- **دراسة الزهراني (2009)** والتي هدفت إلى بناء اختبار محكي المرجع لقياس الكفايات المهنية لدى عينة من معلمي المرحلة الابتدائية بمدينة الرياض، حيث تكونت العينة من (930) معلماً تم اختيارهم بالطريقة العشوائية العنقودية من معلمي المرحلة الابتدائية في الرياض، وطبق عليهم اختبار محكي المرجع يمثل أربع كفايات رئيسة وهي: التخطيط المسبق للدرس، وتنفيذ الدرس، والتقويم، والنمو المهني للمعلم في إلمامه بالخصائص العقلية والاجتماعية لطالب المرحلة الابتدائية تم فيها التحقق من صدق المقياس باستخدام صدق المحتوى بالعرض على الخبراء، وتم التحقق من ثبات الاختبار باستخدام معامل ليفنجستون. كما تم استخدام الإحصاءات الوصفية ومعامل الارتباط ومعامل ثبات ليفنجستون واختبار (ت) للمجموعات المستقلة للتحقق من صدق المقارنة الطرفية. وقد تمتع الاختبار بدرجة مقبولة من الصدق والثبات.
- **كما هدفت دراسة فتيان (2012)** إلى بناء اختبارات محكية المرجع في العلوم العامة للصف الثامن الأساسي في الأردن وفق أسلوب جتمان والتحقق من فاعليتها في ضوء معاملات استرجاعها، وتكونت الاختبارات في صورتها النهائية من عشرة اختبارات جزئية، تم بناء كل اختبار جزئي من ست فقرات مرتبة ترتيباً هرمياً وفق درجة صعوبتها، وأصبح العدد النهائي لفقرات الاختبار كاملاً (60) فقرة، كما تم اختيار عينة عشوائية بسيطة عنقودية

تكونت من (160) طالباً وطالبة من طلبة الصف الثامن الأساسي، وأظهرت النتائج أنَّ معاملات صعوبة فقرات الاختبار متدرجة عبر كل اختبار جزئي، فقد تراوحت بين (0.21- 0.90) أما معاملات التمييز فقد تراوحت بين (0.93-0.21).

ثانياً- دراسات تناولت الكفايات اللغوية:

- **دراسة (Pester. M, 2003)** وقد هدفت إلى معرفة آثار تدخل المهارات اللغوية على اختبار معياري المرجع واختبار محكي المرجع، والتحقق من وجود الاختلاف بين اختبار معياري المرجع كأداة لقياس قبلي، مقابل اختبار محكي المرجع كأداة لقياس بعدي، وتكونت عينة الدراسة من (18) طفلاً متأخراً دراسياً في لغة التخاطب، في جنوب غرب أوهايو، تراوحت أعمارهم بين 25 شهراً و(54) شهراً، واستخدم الباحث اختبارين من إعداد الأخصائيين في أمراض التخاطب أحدهما محكي والآخر معياري، تغطي كل واحدة منها (16) موضوعاً في التخاطب، وتوصلت الدراسة إلى عدة نتائج أهمها: أنَّ الاختبار المحكي المرجع أكثر دقة في تحديد أوجه القصور في التخاطب لدى الأطفال ولا سيما عند استخدامه كأداة للقياس البعدي، كما أنَّ مهارات اللغة المبكرة والكلية كانت مختلفة بشكل ملحوظ بعد (5) أشهر من التدخل العلاجي، لوجود مستويات عجز قبل التدخل.
- **دراسة العواجي (2013)** وقد هدفت إلى التعرف على خصائص طلاب المرحلة الثانوية، والتعرف على أهداف كتاب الكفايات اللغوية، والوقوف على تحقيق كتاب الكفايات اللغوية لأهدافه من خلال تحليل محتواه، وقد توصلت الدراسة إلى أنَّ المقرر لم يراع الفروق الفردية بين الأفراد في الأهداف، فهناك أهداف لا يمكن أنْ يحققها جميع الطلاب، مثل مهارات الدراسة والتعلم الذاتي، حيث بلغت نسبتها في الكتاب الأول (29.35%)، وفي الكتب الثاني (20.9%)، وهذه نسبة مرتفعة أتت على حساب أهداف أخرى، كما دلت نتائج الدراسة على أنَّ بعض الأهداف صعب تحقيقها عن طلاب المرحلة الثانوية وغير ملائمة لهم مثل الهدف الثالث: (تمكين الطالب من تنقيح كتاباته وخطاباته الشفهية المعدة سلفاً من الأخطاء اللغوية والنحوية والإملائية)، وقد أوصت الدراسة بإجراء دراسة لمعرفة مدى إتقان تلاميذ المرحلة الثانوية للكفايات اللغوية.

التعقيب على الدراسات السابقة:

من خلال العرض السابق نجد أنَّ الدراسات السابقة تناولت أوجه مختلفة، وذلك من خلال الأهداف والمنهج والمجتمع والعينة والأداة والنتائج، واستخلص الباحثان من الدراسات السابقة ما يلي:

الهدف: تباينت الدراسات السابقة في أهدافها بعضها هدف إلى بناء اختبارات محكية المرجع معتمدة في ذلك على تحديد النطاق السلوكي للمجال الذي تم بناء الاختبار له كدراسة الزهراني، (2009)، وهناك دراسات أخرى أهتمت بالاختبارات محكية المرجع إلا أنه لم يتم بناء الاختبار على نطاق سلوكي محدد؛ وإنما استخدمت نموذج الاستجابة للمفردة ونظرية القياس الحديثة في بناء وتحليل خصائص الاختبارات المحكية كدراسة (الشريفين، 2006)، والدراسة الحالية تتفق مع الدراسات التي تم تحديد انطقتها السلوكية، حيث قام الباحث بتحديد النطاق السلوكي للكفايات اللغوية لدى طلاب المرحلة الثانوية في مدينة الرياض معتمداً بذلك على المقررات اللغوية التي تم تدريسها، وكذلك الاستعانة بالخبراء والمحكمين في تخصص اللغة العربية والقياس والتقويم.

كما أنَّ الدراسة الحالية تتفق بشكل غير مباشر مع الدراسات التي هدفت إلى التعرف على الكفايات اللغوية لدى طلاب المرحلة الثانوية كدراسة (العواجي، 2013) في حين أنَّ الدراسة الحالية أدق وأخص فقد اقتصرت على بناء اختبار لقياس الكفايات اللغوية (النحوية والإملائية).

بالإضافة إلى أنَّ بعض الدراسات اهتمت بدرجة القطع من حيث أثر كل من نوع المحكم وطول الاختبار على درجة القطع كدراسة (التميمي، 1999)، أما الدراسة الحالية فقد كان تحديد درجة القطع أحد أهدافها والذي تمثل في تحديد درجة القطع المناسبة لكل كفاية لغوية والتي في ضوئها يمكن تصنيف الطلبة إلى متقنين وغير متقنين لهذه الكفاية، حيث استخدم الباحث طريقة ندلسكي لتحديد درجة القطع.

المنهج: اختلفت بعض الدراسات من ناحية المنهج حيث استخدمت بعض الدراسات المنهج السيكومتري والبعض الآخر استخدم منهج بحوث التقويم كدراسة كدراسة (التميمي، 1999) و(الشريفين، 2006) و(الزهراني، 2009) و(فتيان، 2012) و(العواجي، 2013) وهو المنهج الذي تم استخدامه في الدراسة الحالية؛ كونه المنهج المناسب في مثل هذا النوع من الدراسات التي تهدف إلى بناء اختبارات محكية المرجع.

المجتمع والعينة: اختلفت الدراسات السابقة فيما بينها من حيث نوعية عيناتها، فقد تم تطبيق بعض الدراسات على المعلمين كدراسة (الزهراني، 2009) وبعض الدراسات تكونت عينتها من طلبة الجامعات كدراسة (الشريفين، 2006) وكما أنَّ بعض الدراسات تم تطبيقها على طلبة المرحلة الابتدائية والمتوسطة كدراسة (التميمي، 1999) و(فتيان، 2012) في حين اتفقت الدراسة الحالية في كون تطبيق العينة على طلاب المرحلة الثانوية كدراسة (العواجي، 2013).

الأداة: اتفقت الدراسات السابقة من حيث الأدوات المستخدمة فيها، حيث استخدمت الاختبار من نوع الاختيار من متعدد كدراسة (التميمي، 1999) و(الشريفين، 2006) و(الزهراني، 2009) و(فتيان، 2012) و(العواجي، 2013) وهي الأداة نفسها التي استخدمها الباحث بالدراسة الحالية حيث قام ببناء اختبار محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض حيث تكون في صورته النهائية من (40) فقرة من نوع الاختيار من متعدد رباعي البدائل.

3- منهجية البحث وإجراءاته

منهج الدراسة

استخدم الباحثان منهج بحوث التقويم كونه المنهج المناسب في مثل هذا النوع من الدراسات التي تهدف إلى بناء اختبارات محكية المرجع.

أداة البحث:

في ضوء أهداف الدراسة قام الباحث ببناء اختبار تحصيلي محكي المرجع لقياس الكفايات اللغوية لدى طلاب المرحلة الثانوية، وقد قام الباحث باتباع الخطوات المتبعة لبناء هذا النوع من الاختبارات كما أشارت لذلك الدراسات السابقة مثل دراسة (علام، 2001، أ) و(البناء، 2011) و(العنزي، 2003) و(الطراونه، 2006) و(ناجي، 2010) و(الزهراني، 2009)

الخصائص السيكومترية لأداة الدراسة:

أولاً- الصدق الظاهري (صدق المحكمين):

تم عرض الأداة على مجموعة من المحكمين للحكم على مدى اتفاق البند مع الهدف وقياسه ومدى تمثيله لذلك الهدف، وبلغ مجموع عدد المحكمين (13) فردًا من أعضاء هيئة التدريس بالجامعات السعودية ومن منتسبي

وزارة التعليم، حيث طُلب منهم دراسة الأداة وإبداء رأيهم فيها من حيث مدى مناسبة الفقرة للمحتوى، كما طُلب منهم النظر إلى مدى شمولية الأداة وتنوع محتواها وتقويم الصياغة اللغوية والإخراج، وأية ملاحظات يرونها مناسبة فيما يتعلق بالتعديل أو التغيير أو الحذف وفق ما يراه المحكم ضرورياً، وقد قام الباحث بدراسة ملاحظات المحكمين واقتراحاتهم والتعديل عليها في ضوء توصيات وآراء المحكمين.

وقد اعتبر الباحثان الأخذ بملاحظات المحكمين، وإجراء التعديلات المشار إليها أعلاه بمثابة الصدق الظاهري، وصدق المحتوى لأداة الدراسة، ولذا اعتبر أنَّ الأداة صالحة لقياس ما وضع له.

ثانياً- صدق الاتساق الداخلي:

تم التحقق من صدق الاختبار بصورته الأولية من خلال صدق الاتساق الداخلي وللتحقق من صدق الاختبار تم استخراج معامل ارتباط بوينت باسيريال بين الفقرة والدرجة الكلية، ومعاملات الارتباط هذه تعتبر دليلًا يراه علماء القياس على صدق محتوى الاختبار ومنهم (أبو حطب وآخرون، 2008)، كما تعتبر من أدق الطرق المعروفة لحساب الصدق وهي تؤدي إلى معرفة معامل الصدق بصورة صحيحة، جدول (1) التالي يوضح معاملات الارتباط.

الجدول (1) معاملات ارتباط بوينت بايسيريال بين الفقرة والدرجة الكلية

الفقرة	معامل ارتباط بوينت بايسيريال	الفقرة	معامل ارتباط بوينت بايسيريال
1	42.0**	24	0.56**
2	43.0**	25	0.57**
3	0.42*	26	0.42**
4	0.13	27	0.42*
5	0.32*	28	0.55**
6	0.35*	29	0.65**
7	0.34*	30	0.57**
8	0.37*	31	0.62**
9	0.50**	32	0.47**
10	0.37*	33	0.56**
11	0.51**	34	0.32*
12	0.52**	35	0.58**
13	0.63**	36	0.58**
14	0.58**	37	0.40*
15	0.22	38	0.31*
16	0.62**	39	0.33*
17	0.36*	40	0.23
18	0.68**	41	0.42**
19	0.22	42	0.39*
20	0.60**	43	0.19
21	0.42*	44	0.37*
22	0.40*	45	0.38*
23	0.41*		

**دال احصائياً عند مستوى دلالة $\alpha \leq 0.01$ * دال احصائياً عند مستوى دلالة $\alpha \leq 0.05$

تبين من خلال الجدول (1) تراوح معاملات الارتباط من (0.19- 0.68) حيث هناك فقرات لها ارتباط بالدرجة الكلية ودالة احصائياً، وفقرات لم يكن لها دلالة احصائية حيث سيتم حذف الفقرات غير الدالة احصائياً.

ثانياً- الثبات:

للتحقق من ثبات المقياس تم استخراج معامل الثبات بطريقتين:

كودر ريتشاردسون (20) وذلك لمناسبته لفقرات التي اجابتها (0، 1)، كما تم استخدام معامل التجزئة النصفية. وجدول (2) يوضح ذلك.

جدول (2) نتائج معاملات الثبات للاختبار بصورته الأولية

عدد الفقرات	معامل الثبات (كودر- رتشارسون)	معامل الثبات التجزئة النصفية
45	0.86	0.75

تبين من خلال جدول(2) أنَّ معامل الثبات بطريقة كودر –ريتشاردسون بلغت (0.86) وهو معامل ثبات مرتفع، كما بلغ معامل الثبات بطريقة التجزئة النصفية (0.75) وهو معامل ثبات جيد، مما سبق يدل على أنَّ المقياس يتصف بالثبات.

الأساليب الإحصائية المستخدمة لمعالجة البيانات:

استخدم الباحثان الأساليب الإحصائية التالية:

1- الإحصاءات الوصفية: ومنها المتوسط الحسابي والانحراف المعياري والنسبة المئوية.
2- معامل ارتباط (بوينت بايسيريال).
3- معامل ثبات (ألفا كرونباخ)، ومعامل ثبات ليفنجستون الخاص بالاختبار محكي المرجع.
4- اختبار (ت) للمجموعات المستقلة للتحقق من صدق المقارنات الطرفية.

4- عرض النتائج وتفسيرها

- **نتائج الإجابة عن التساؤل الأول: هل يتوافر للاختبار محكي المرجع الذي تم بناؤه والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض مؤشر مقبول من الصدق؟**

وللإجابة عن التساؤل الأول تم التحقق من أنواع الصدق التالية:

أولا صدق المحتوى: حيث تم عرض الاختبار على مجموعة من المحكمين الخبراء في تخصص اللغة العربية والقياس والتقويم، وذلك للتحقق من مدى اتفاق كل مفردة اختبارية مع هدفها السلوكي.

- صدق المفهوم: تمّ حساب معاملات ارتباط بوينت بايسيريال للعلاقة بين المفردة والدرجة الكلية للاختبار ويبين الجدول رقم (3) معاملات الارتباط لكل مفردة من مفردات الاختبار.

جدول (3): نتائج معاملات ارتباط الفقرات والدرجة الكلية ن=983

المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط
1	.25**0	11	.42**0	21	.37**0	31	.45**0
2	.31**0	12	0.36**	22	.22**0	32	.37**0
3	.24**0	13	.29**0	23	.30**0	33	.26**0

المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط
4	.22**0	14	.33**0	24	.34**0	34	.36**0
5	.31**0	15	.36**0	25	.43**0	35	.27**0
6	.32**0	16	.28**0	26	.42**0	36	.38**0
7	.31**0	17	.28**0	27	.41**0	37	.31**0
8	.22**0	18	.37**0	29	.40**0	38	.28**0
9	.31**0	19	.41**0	29	.29**0	39	.31**0
10	.26**0	20	.24**0	30	.49**0	40	.46**0

** دال احصائياً عند مستوى دلالة $\alpha \leq 0.05$

يلاحظ من الجدول (3) أنَّ قيم معاملات الارتباط لمُفردات الاختبار تراوحت من 0.22 وحتى 0.49 مع الدرجة الكلية وهي معاملات دالة احصائياً عند مستوى دلالة $\alpha \leq 0.05$ مما يدل على تحقق دلالات صدق المفهوم للاختبار.

- ما دلالة الصدق التمييزي بدلالة المقارنة الطرفية لاختبار كفايات اللغة العربية؟

تم التوصل إلى دلالات الصدق التمييزي للاختبار باستخدام اختبار t للعينتين المستقلتين وفقًا للدرجة (مرتفعة، منخفضة)، وذلك كما هو مبين في جدول (4) التالي.

جدول (4) نتائج اختبار (ت) لاختبار الكفايات في اللغة العربية

الفئة				اختبار(ت) لتكافؤ الأوساط الحسابية		
عليا (235)		دنيا (325)				
الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	الانحراف المعياري	النسبة الحرجة (ت)	درجة الحرية	احتمالية الخطأ
31.26	3.15	18.82	3.22	49.78	648	0.00*

*دال احصائياً عند مستوى دلالة $\alpha = 0.05$

يتضح من الجدول (4) وجود فرق جوهري عند مستوى الدلالة (α=0.05) بين الوسطين الحسابيين للاختبار يعزى للفئة (عليا ودنيا)؛ لصالح العليا مقارنة بالدنيا؛ حيث بلغ المتوسط الحسابي للفئة الدنيا(18.82) بانحراف معياري (3.22) والمتوسط الحسابي للفئة العليا (31.26) بانحراف معياري (3.15) وبلغت قيمة اختبار (ت=49.78) وهي قيمة دالة احصائياً عند مستوى دلالة α =0.05. مما يشير إلى تحقق الصدق التمييزي للاختبار. من خلال النتائج السابقة يتبين أنَّ المقياس يتمتع بدلالات صدق مرتفعة من خلال صدق التكويني (صدق المحتوى وصدق الاتساق الداخلي، والصدق التمييزي). كما أظهرت النتائج وجود درجة كبيرة من الصدق الداخلي.

- **نتائج الإجابة عن التساؤل الثاني: هل يتوافر للاختبار محكي المرجع الذي تم بناؤه والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض درجة مقبولة من الثبات؟**

وللإجابة عن سؤال الدراسة الثاني؛ فقد تم حساب معامل ثبات عن طريق إعادة التطبيق، كما تم حساب معامل الثبات بطريقة ألفا-كرونباخ لمُفردات الاختبار، كما تم استخدام معامل الثبات بطريقة التجزئة النصفية وذلك على النحو الآتي:

1- ما معامل ثبات اختبار كفايات اللغة العربية بطريقة إعادة التطبيق؟

للإجابة عن سؤال الدراسة تم استخراج معامل الثبات بطريقة إعادة التطبيق حيث قام الباحث بتطبيق الاختبار على عينة الدراسة من ثم بعد أسبوعين قام الباحث بإعادة التطبيق ومن ثم تم إيجاد معامل الارتباط بين الدرجتين وهو معامل الثبات بإعادة التطبيق. وجدول (5) يبين ذلك.

الجدول (5) نتائج معامل الثبات بطريقة إعادة التطبيق

الاختبار	العينة	معامل الارتباط	الدلالة الاحصائية
اختبار كفايات اللغة العربية	30	0.701	0.00*

يتبين من خلال جدول(5) أنَّ معامل الارتباط بين التطبيق الأول للاختبار والتطبيق الثاني بلغ (0.701) وهو معامل ارتباط مرتفع ويدل على ثبات الاختبار بطريقة إعادة التطبيق.

ما معامل ثبات اختبار كفايات اللغة العربية بطريقة ألفا-كرونباخ؟

وللإجابة عن سؤال الدراسة تم استخراج معامل الثبات بطريقة الفا كرنباخ. وجدول (6) التالي يبين ذلك. وتعتمد هذه الطريقة على مدى تجانس درجات الاختبار المستخدم، وهي تكشف عن تشتت درجات التوزيع الممثل بدرجات المتعلمين.

جدول (6) نتائج معامل الثبات بطريقة معامل ألفا كرونباخ

الاختبار	العينة	معامل الثبات ألفا كرونباخ
اختبار كفايات اللغة العربية	983	0.842

يتبين من خلال جدول(6) أنَّ معامل الثبات بطريقة ألفا كرونباخ بلغ (0.842) وهو معامل ارتباط مرتفع ويدل على ثبات الاختبار بطريقة ألفا كرونباخ.

- **نتائج الإجابة عن التساؤل الثالث: هل يتوافر للاختبار محكي المرجع الذي تم بناؤه والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض درجة مقبولة من التمييز؟**

للإجابة عن سؤال الدراسة الثالث تم حساب معامل التمييز لفقرات الاختبار باستخدام المعادلة:

$$\text{معامل التمييز} = \frac{\text{س}-\text{ص}}{\text{ن}}$$

حيث:

س: عدد طلاب الفئة العليا في التحصيل الذين أجابوا على السؤال إجابة صحيحة.

ص: عدد طلاب الفئة الدنيا في التحصيل ممن أجابوا على السؤال إجابة صحيحة.

ن: عدد أفراد إحدى المجموعتين.

وجدول (7) التالي يبين نتائج معاملات تمييز الاختبار.

جدول (7) نتائج معاملات تمييز الاختبار ن=983

المفردة	معامل التمييز	المفردة	معامل التمييز	المفردة	معامل التمييز	المفردة	معامل التمييز
1	0.27	11	0.46	21	0.38	31	0.52
2	0.38	12	0.21	22	0.26	32	0.44
3	0.24	13	0.36	23	0.34	33	0.28
4	0.24	14	0.38	24	0.38	34	0.41

المفردة	معامل التمييز	المفردة	معامل التمييز	المفردة	معامل التمييز	المفردة	معامل التمييز
5	0.38	15	0.43	25	0.48	35	0.35
6	0.39	16	0.36	26	0.48	36	0.45
7	0.31	17	0.34	27	0.43	37	0.36
8	0.30	18	0.40	28	0.47	38	0.33
9	0.20	19	0.41	29	0.33	39	0.32
10	0.32	20	0.23	30	0.58	40	0.34

يتبين من خلال جدول (7) أنَّ معاملات التمييز للاختبار تراوحت ما بين (0.20 – 0.58) وهي معاملات تمييز مقبولة إلى جيدة التمييز. وكان أعلى معامل تمييز للمفردة (30) وأدنى معامل تمييز للمفردة (9).

- **نتائج الإجابة عن التساؤل الرابع: هل يتوفر للاختبار محكي المرجع الذي تم بناؤه والمستخدم في قياس الكفايات اللغوية لدى طلاب المرحلة الثانوية بمدينة الرياض درجة مقبولة من الصعوبة ؟**

وهو عبارة عن النسبة المئوية من الطلاب الذين أجابوا عن السؤال إجابة صحيحة ويحسب بتطبيق المعادلة التالية:

$$\text{معامل صعوبة السؤال} = \frac{\text{س}}{\text{ن}} \times 100$$

حيث س: عدد الطلاب الذين أجابوا على السؤال إجابة صحيحة. ن: مجموع الطلاب

وجدول (8) التالي يوضح ذلك.

جدول (8) نتائج معاملات الصعوبة للاختبار ن=983

المفردة	معامل الصعوبة	المفردة	معامل الصعوبة	المفردة	معامل الصعوبة	المفردة	معامل الصعوبة
1	0.69	11	0.74	21	0.64	31	0.69
2	0.62	12	0.62	22	0.64	32	0.60
3	0.69	13	0.57	23	0.63	33	0.65
4	0.58	14	0.59	24	0.61	34	0.65
5	0.68	15	0.59	25	0.69	35	0.66
6	0.69	16	0.58	26	0.66	36	0.61
7	0.64	17	0.63	27	0.73	37	0.58
8	0.64	18	0.75	28	0.69	38	0.64
9	0.63	19	0.70	29	0.63	39	0.50
10	0.57	20	0.66	30	0.62	40	0.64

يتبين من خلال جدول(8) أنَّ معاملات الصعوبة للاختبار تراوحت ما بين (0.50- 0.75) وهي معاملات صعوبة جيدة. وكان أعلى معامل صعوبة للمفردة (18) وأقل معامل صعوبة للمفردة (39). وبشكل عام أفضل معامل صعوبة للسؤال أو المفردة هو 50% وما حولها.

مناقشة النتائج:

اتفقت نتائج بعض الدراسات وجود درجة كبيرة من الصدق الداخلي والثبات كدراسة و(الشريفين، 2006) و(الزهراني، 2009) و(فتيان، 2012) و(العواجي، 2013).

كما اتفقت بعض النتائج في استخدام معامل التمييز كدراسة (الشريفين، 2006) واختلفت مع دراسة و(الزهراني، 2009) و(فتيان، 2012) و(العواجي، 2013). وأيضاً اتفقت بعض نتائج الدراسات في استخدام معامل الصعوبة كدراسة (الشريفين، 2006)، واختلفت مع دراسة (الزهراني، 2009) و(فتيان، 2012) و(العواجي، 2013).

واتفقت كذلك بعض النتائج كدراسة (التميمي، 1999) و(الزهراني، 2009) و(ناجي، 2010) في استخدام درجة القطع في أبحاثهم (ندلسكي أو أنجوف) واختلفت مع دراسة (الشريفين، 2006) و(فتيان، 2012) و(العواجي، 2013).

توصيات البحث ومقترحاته

1. تطبيق الاختبار الذي تم بناؤه في الدراسة الحالية على الطلبة؛ لتشخيص مدى تمكنهم في الكفايات اللغوية.
2. تطبيق دراسة مماثلة لبناء اختبار محكي المرجع لقياس الكفايات اللغوية لدى طلبة الابتدائية والمتوسطة
3. توصية خبراء القياس والتقويم بإيجاد طرق جديدة لبناء الاختبارات محكية المرجع في كافة مواضيع التربية.
4. إجراء دراسة لإيجاد استراتيجية جديدة لبناء الاختبارات محكية المرجع قائمة على طرق إحصائية تحليلية، عوضاً عن الطريقة السائدة المعتمدة على تحكيم الخبراء.
5. إجراء دراسة لبناء بنك أسئلة لمقرر كفايات اللغة العربية للمرحلة الثانوية.
6. إجراء دراسة مماثلة لبناء اختبار محكي المرجع لقياس الكفايات اللغوية لدى الطلبة في بيئات محلية وعربية.
7. إجراء دراسة مماثلة لبناء اختبار محكي المرجع لقياس الكفاية القرائية لدى طلاب المرحلة الثانوية.

قائمة المراجع

أولاً- المراجع بالعربية:

- أبو حطب، فؤاد وعثمان، سيد وصادق، آمال (2008). التقويم النفسي. ط(4). القاهرة: مكتبة الانجلو المصرية.
- أحمد، مرشود عيسى (2009). العوامل المؤثرة في ضعف المستوى وتدني التحصيل الدراسي في اللغة العربية لطلاب الأول والثاني الثانوي بولاية القضارف. رسالة ماجستير ، جامعة السودان للعلوم والتكنولوجيا، الخرطوم.
- البناء، مأمون علي (2011). بناء اختبار محكي المرجع لقياس الكفايات الإحصائية لدى طلبة الدراسات العليا بكليات التربية في الجامعات اليمنية، رسالة ماجستير غير منشورة، جامعة الملك سعود، الرياض.
- التميمي، خالد حسن (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 غير منشورة، قسم علم النفس، كلية التربية، جامعة أم القرى، مكة المكرمة.
- الحربي، خليل عبدالرحمن (2013). مستوى أداء خريجي التعليم الثانوي في المملكة العربية السعودية. مجلة رسالة التربية وعلم النفس، الجمعية السعودية للعلوم التربوية والنفسية (جستن)، (41)، 125-144.

- الحموري، هند عبدالمجيد؛ والكحلوت، أحمد إسماعيل (2009). استراتيجية مقترحة لبناء اختبارات تشخيصية محكية المرجع. مجلة العلوم التربوية والنفسية، كلية التربية، جامعة البحرين، 9(3)، ص ص153-175.
- الزهراني، سامي صالح (2009). بناء اختبار محكي المرجع لقياس الكفايات المهنية لدى عينة من معلمي المرحلة الابتدائية بمدينة الرياض. ماجستير غير منشورة، قسم علم النفس، كلية التربية، جامعة الملك سعود، الرياض.
- الشريفين، نضال كمال (2006). الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق النظرية الحديثة في القياس التربوي. مجلة العلوم التربوية والنفسية، م(7)، ع(4)، ص80.
- الصبحي، محمد علي (2001). بناء اختبار محكي المرجع لقياس الكفايات الرياضية في المفاهيم الهندسية للمرحلة الابتدائية في مدارس مكة المكرمة الحكومية. رسالة ماجستير غير منشورة، كلية التربية، جامعة أم القرى، مكة.
- الطراونه، عيسى عبدالوهاب (2006). بناء اختبار محكي المرجع لقياس كفايات المعلمين في بناء اختبارات التحصيل. ماجستير غير منشورة، قسم الإرشاد والتربية الخاصة، كلية العلوم التربوية. جامعة مؤتة، الأردن.
- الطريري، عبدالرحمن سليمان (2009). القياس النفسي والتربوي نظريته- أسسه- تطبيقاته، ط(2)، الرياض: مكتبة الرشد.
- علام، صلاح الدين محمود (2001، أ). الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية. ط(2). القاهرة: دار الفكر العربي.
- علام، صلاح الدين محمود (2000). تحليل بيانات البحوث النفسية والتربوية والاجتماعية. القاهرة: دار الفكر العربي.
- علام، صلاح الدين محمود (2001، ب). القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة: دار الفكر العربي.
- العنزي، حسين عزيز (2003). بناء اختبار تحصيلي لقياس درجة إتقان المهارات الأساسية في مادة الرياضيات لدى تلاميذ الصفوف العليا في المرحلة الابتدائية. رسالة ماجستير ، كلية التربية، جامعة الملك سعود، الرياض.
- العواجي، خالد سعيد (2013). تقويم كتاب الكفايات اللغوية للصف الأول الثانوي في ضوء أهدافه، رسالة ماجستير غير منشورة، الجامعة الإسلامية بالمدينة المنورة.
- عوض، أحمد (2010). فعالية استراتيجية مقترحة في علاج الضعف القرائي والكتابي والتحصيلي في اللغة العربية لدى بعض تلاميذ الصف التاسع من التعليم الأساسي، مجلة كلية التربية، المنصورة.
- فتيان، أمل محمد (2012). بناء اختبارات تشخيصية محكية المرجع في العلوم العامة للصف الثامن الأساسي في الأردن وفق أسلوب جتمان والتحقق من فاعليتها في ضوء معاملات استرجاعها. رسالة ماجستير ، جامعة مؤتة.
- مجيد، سوسن شاكر (2007). أسس بناء الاختبارات والمقاييس النفسية والتربوية. عمان: دار ديبونو للطباعة والنشر والتوزيع.
- مدكور، علي أحمد (1991). تدريس فنون اللغة العربية، القاهرة: مكتبة الفلاح للنشر والتوزيع.
- ناجي، حازم يحيى (2010). بناء اختبار تشخيصي محكي المرجع في المفهومات الأساسية لعلم النفس التربوي. رسالة ماجستير غير منشورة، قسم علم النفس، جامعة صنعاء، اليمن.

ثانياً- المراجع بالإنجليزية References

- Pester. M.(2003). Language Intervention Effects on Norm referenced. Criterion referenced test Scores. Submitted in the Partial Fulfillment of The Requirements for The degree of Master of Arts. Department of Speech Pathology and Audiology.