Journal of Educational and Psychological Sciences
Volume (3), Issue (25) : 30 Oct 2019
P: 26 - 48

AJSRP
ISSN: 2522-3399

مجلة العلوم التربوية والنفسية
المجلد (3)، العدد (25) : 30 أكتوبر 2019 م
ص: 26 - 48

Evaluation of School Tests Prepared by Science Teachers at the Intermediate Stage in the City of Tabuk in Terms of Planning, Formatting and Correcting Them

SMAH SULIMAN S. ALBALAWI **SHAHER KHALED SULIMAN**

Faculty of Education &Arts || University of Tabuk || KSA

Abstract: This research aimed to identify the commitment degree of the science teachers in the middle stage to the criteria for planning, formatting, and correcting school tests. To achieve the objective of the research, the researcher used descriptive analytical method to identify the competencies and skills of teachers in the evaluation of school tests. The study was carried out on a sample of (100) teachers of science, were randomly selected, in the intermediate stage in Tabuk region. A questionnaire consisted of (21) articles was distributed on them discussing the planning and correction of school tests. The questionnaire was applied to a pilot sample of (30) teachers, and presented to a group of arbitrators; to ascertain its sincerity and persistence. Also, thirty (30) tests sheets were analyzed using a content analysis form consists of (29) articles discuss the technical standards and the test paragraphs' formatting school tests preparation process; it was presented to a group of arbitrators to verify its validity and consistency. The results of the research came as follows, That the degree of science teachers' commitment in planning and correcting school tests obtained an average of (4.73- 3.38), i.e., it is high, and tests formatting skills have obtained an average of (80%) which means that it is intermediate. The results also showed that there were no statistically significant differences at the level of (α0.05) due to the qualification variable. There were statistically significant differences at the level of indication (α0.05) due to the variable of experience for those with more than ten years' experience. The study issued a set of recommendations, including: the need to hold training courses in the construction of the achievement tests for teachers, and encourage teachers to take advantage of each other's experiences and exchange experiences among them, and urged teachers to arbitrate the tests after written by their colleagues experienced teachers to ensure their suitability.

Keywords: Evaluation, School Tests, Planning, Formatting, Correcting.

تقويم الاختبارات المدرسية التي تعدها معلمات العلوم في المرحلة المتوسطة بمدينة تبوك من حيث تخطيطها، وإخراجها، وتصحيحها

سماح سليمان صويلح البلوي **شاهر خالد سليمان**

كلية التربية والآداب || جامعة تبوك || المملكة العربية السعودية

الملخص: هدف هذا البحث إلى التعرف على درجة التزام معلمات العلوم في المرحلة المتوسطة بمعايير التخطيط للاختبار وإخراجه وتصحيحه، والكشف عن مدى استيفاء اختبارات معلمات المرحلة المتوسطة لمعايير بناء فقرات الاختبار. ولتحقيق أهداف البحث، استخدم الباحثان المنهج الوصفي التحليلي للوقوف عند كفايات ومهارات المعلمات في تقويم الاختبارات المدرسية. وقد تم تطبيق الدراسة على عينة بلغت (100) معلمة من معلمات مادة العلوم في المرحلة المتوسطة بمنطقة تبوك تم اختيارهن بطريقة عشوائية، وتم توزيع استبانة عليهن مكونة من (21) فقرة تناقش محوري تخطيط وتصحيح الاختبارات المدرسية، وقد تم تطبيق هذه الاستبانة على عينة استطلاعية قوامها (30) معلمة، وتم عرضها على مجموعة من المحكمين؛ للتأكد من صدقها وثباتها. وقد جرى أيضاً تحليل (30)

اختباراً من أوراق الاختبارات المعدة من قبل المعلمات؛ وذلك باستخدام استمارة تحليل المحتوى، والتي بلغ عدد فقراتها (29) فقرة تناقش محوري الاهتمام بالمعايير الفنية وبناء الفقرات في إعداد الاختبارات المدرسية؛ وتوصل البحث إلى بعض النتائج من أهمها: إن درجة التزام معلمات العلوم بمهارات تخطيط وتصحيح الاختبارات التحصيلية حصلت على متوسط (4.73- 3.38) أي أنها عالية، أما مهارات إخراج الاختبار فقد حصلت على ما نسبته (80%) وهو ما يعني أنها متوسطة. كما أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$) تُعزى لمتغير المؤهل، ووجود فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$) تُعزى لمتغير الخبرة لصالح من يمتلكن خبرة تتجاوز العشرة سنوات. وخرج البحث بمجموعة من التوصيات منها: ضرورة عقد دورات تدريبية في بناء الاختبارات التحصيلية للمعلمات، وتشجيع المعلمات على الاستفادة من خبرات بعضهن البعض وتبادل الخبرات فيما بينهن، وحث المعلمات على تحكيم الاختبارات بعد كتابتها من قبل زميلاتهن المعلمات ذوات الخبرة للتأكد من ملاءمتها.

الكلمات المفتاحية: التقويم، الاختبارات المدرسية، التخطيط، الإخراج، التصحيح.

1- مقدمة البحث

يخطئ البعض حين يظن أن عمل المعلم ما هو إلا عمل روتيني يتلخص في إلقاء بعض الدروس داخل جدران حجرة الدراسة وحسب، فالمعلم المعاصر يمارس ألواناً من الأنشطة العملية الدقيقة؛ فيستطيع أن يصمم الاختبارات، ويضع الأسئلة المناسبة لقياس المستويات المختلفة، ويحدد الاختبارات المناسبة لقياس الجوانب المختلفة للتعلم، سواء كانت قدرات عقلية أو سمات شخصية. كما بوسعه استخدام الوسائل والطرق المختلفة لتحليل ما يحصل عليه من نتائج خاصة بتحصيل التلاميذ ولكي يحدث ذلك، فلابد من إلمام المعلم بمعايير الاختبارات والأساليب المختلفة لتصنيف الاختبارات ومعرفة شروط الاختبار الجيد (فتح الله، 2016: 280).

وتعد الاختبارات التحصيلية واحدة من أهم الأدوات المستخدمة في عملية قياس النتاجات التعليمية؛ للوقوف على مدى تحقق الأهداف التربوية والتعليمية، ويتوقف على نتائجها اتخاذ قرارات غاية في الأهمية سواء بالنسبة إلى طرائق التعليم، أو المناهج الدراسية، أو المصير الذي سيؤول إليه المتعلمون (عمارة، 2008).

مشكلة البحث

تشكِّل عملية التقويم أحد العناصر الأساسية للعملية التربوية والتعليمية، وتعدُّ الاختبارات جزءاً لا يتجزأ من عملية التقويم. كما تُعدّ من الوسائل المهمّة التي يُعتمد عليها في قياس وتقييم قدرات الطلبة على اختلاف مستوياتهم الدراسية، فمن خلالها يمكن الكشف عن عناصر النجاح، وتشخيص نقاط القوة والضعف لدى المتعلمين من أجل الارتقاء وتطوير مستواهم الأكاديمي. ومن هذا المنطلق فقد حرص التربويون على أن تكون هذه الاختبارات ذات كفاءة عالية، وتتصف بالموضوعية والصدق والثبات وحسن الإعداد والإخراج، وبالشكل الذي تكون فيه عوناً للمعلم والمتعلم على حد سواء؛ لأنها إذا لم تعد وفق معايير الاختبار الجيد، فإنها ستلحق الأذى والضرر بالأفراد ذوي العلاقة نفسياً وتربوياً واجتماعياً (العون، 2016) و(مراد وسليمان، 2005).

ويعد إغفال معايير إعداد الاختبارات المدرسية أحد أهم الأسباب التي تؤدي إلى عدم تحقيق هذه الاختبارات للغايات التي أُعدت من أجلها، وتؤكد العديد من الدراسات على وجود مستوى متوسطٍ إلى عالٍ في إغفال بعض معايير إعداد وتقويم الاختبارات المدرسية؛ إذ تؤكد دراسة المالكي (2017) على أنَّ القصور الحاصل في معايير إعداد الاختبارات المدرسية في محاور تعليمات الاختبار، وعدم تناسب الاختبار مع الوقت الزمني المحدد لأجله، والدرجة المعطاة لكل سؤال، هي من أهم الأسباب التي تؤدي إلى فشل الاختبارات في تحقيق أهداف، وعاملٌ محوري في تراجع معدلات التحصيل الأكاديمي عند العديد من الطلاب.

كما وتؤكد دراسة الرويلي (2016) على أن هناك تدنٍ ملحوظ في مستوى معرفة الطلاب المعلمين بمهارات القياس والتقويم الواجب اتباعها في عملية بناء الاختبارات المدرسية، وهو ما ينعكس بشكلٍ سلبي على نجاعة الاختبارات المدرسية وقدرتها على تحقيق الأهداف المرجوة منها؛ وهو ما ينعكس بدوره على ضعف النتاج الأكاديمي والتحصيلي للطلاب.

أسئلة البحث:

في ضوء المشكلة البحثية، يسعى هذا البحث إلى الإجابة عن التساؤلات التالية:

1- ما درجة التزام معلمات العلوم للمرحلة المتوسطة بمهارات التخطيط للاختبارات وفق معايير تخطيط للاختبار؟
2- ما درجة التزام معلمات العلوم للمرحلة المتوسطة بمعايير تصحيح الاختبار؟
3- ما مدى التزام معلمات المرحلة المتوسطة بمعايير إخراج الاختبار؟
4- هل توجد فروق ذات دلالة إحصائية عند مستوى ($\alpha \leq 0.05$) في درجة التزام معلمات المرحلة المتوسطة بمعايير تخطيط وتصحيح الاختبار تُعزى إلى متغيري: الخبرة، والمؤهل العلمي والتفاعل بينهما؟

أهداف البحث

يسعى البحث الحالي إلى تحقيق الأهداف التالية:

1- التعرف على درجة التزام معلمات المرحلة المتوسطة في مدارس منطقة تبوك بمعايير تخطيط وإخراج وتصحيح الاختبار الجيد.
2- الكشف عن مدى استيفاء اختبارات معلمات المرحلة المتوسطة لمعايير بناء فقرات الاختبار.
3- فحص مدى تأثير متغيرات (الخبرة، والمؤهل العلمي) عند مستوى الدلالة ($\alpha \leq 0.05$) بين إجابات العينة؛ حول معايير تخطيط وتصحيح وإخراج الاختبار التحصيلي الجيد.
4- المقارنة بين المعلمات وفقاً لعدد من متغيرات، الخبرة، والمؤهل العلمي والتفاعل بينهم في مدى توافق اختباراتهم وفقاً لمعايير تخطيط وتصحيح واخراج الاختبار التحصيلي الجيد.

أهمية البحث

تتمثل أهمية البحث فيما يلي:

- الأهمية النظرية:
  - ندرة الدراسات والأبحاث التي تناقش موضوع تقويم الاختبارات المدرسية في المرحلة المتوسطة، لاسيما في البيئة السعودية.
  - توجيه الباحثين إلى تبني توجهات جديدة في أبحاثهم العلمية، ومساعدة الباحثين المستقبليين العازمين على تناول موضوع البحث.
- الأهمية التطبيقية:
  - قد تفيد نتائج البحث في تزويد الجهات المسؤولة في وزارة التربية والتعليم بمؤشرات حول مدى جودة الاختبارات التي تقوم المعلمات بإعدادها ومدى مراعاتها لشروط الاختبار الجيد.
  - تسليط الضوء على آليات إعداد وتخطيط وإخراج وتصحيح الاختبارات المدرسية وفق الأساليب التي تضمن تحقيق الغايات التي أُعدت من أجلها هذه الاختبارات.

- لفت انتباه المسؤولين والمختصين إلى ضرورة العمل على عقد الدورات التدريبية واختبارات الكفايات التدريسية للمعلمات والمعلمين؛ لقياس وتطوير مستوى إلمامهم بمعايير إعداد الاختبارات المدرسية.

حدود البحث

- **الحدود الموضوعية:** اقتصر البحث على معرفة درجة التزام المعلمات بمهارات بناء الاختبارات من حيث التخطيط للاختبارات، وإخراجها من حيث الشكل والمضمون، وتصحيحها.
- **الحدود البشرية:** تم تطبيق الدراسة على (100) معلمة من معلمات العلوم للمرحلة المتوسطة في منطقة تبوك والبالغ عددهن (138) معلمة والموزعات على (58) مدرسة في منطقة تبوك.
- **الحدود المكانية:** تم إجراء البحث في محافظة تبوك بالمملكة العربية السعودية.
- **الحدود الزمانية:** تم إجراء البحث في العام الدراسي 1438 /1439هـ.

مصطلحات البحث

- **التقويم:**

التقويم لغةً: التقويم على زنة التفعيل مشتق من الفعل الثلاثي من مادة (قَوَمَ) قد جاء في لسان العرب بمعنى التقدير، أقام السلعة وقومها أي قدرها وأقمت الشيء وقومته فقام بمعنى الاستقامة واعتدال الشيء واستوائه، وقومت الشيء فهو قويم أي مستقيم (ابن منظور، 1993: 431). وقيل قوم الشيء أي عدَّله وأزال العوج عنه، وقيم الشيء تقييماً أي قدر قيمته، وقوم درأه أي أزال اعوجاجه، أو جعل له قيمة معلومة. (مجمع اللغة العربية، 2010: 768).

التقويم اصطلاحاً: يعرف على أنه: "مجموعة من الإجراءات التي يتم خلالها جمع بيانات حول فرد أو مشروع أو ظاهرة، ودراسة هذه البيانات بأسلوب علمي للتأكد من مدى تحقيق أهدافاً أُعدت سابقاً من أجل اتخاذ قرارات معينة"(مصبح، 2013: 18).

كما ويعرف أيضاً بأنه: "طريقة يتم فيها اتخاذ قرار حول قيمة شيء ما، بناءً على بعض البيانات المنظمة" (Martorella, 1991: 319).

ويعرفه الباحثان إجرائياً بأنه: عملية يتم من خلالها إصدار الحكم على مدى وصول الاختبارات المدرسية إلى أهدافها، ومدى تحقيقها لأغراضها، والكشف عن نواحي النقص فيها، ومن ثم اقتراح الآليات والوسائل المناسبة لسد هذا النقص ومعالجته.

- **الاختبارات المدرسية:**

تعرف بأنها: "الاختبارات التي تقدم في جميع نظم الامتحانات بالمؤسسات المدرسية وهي اختبارات أداء أقصى، أي أنها تقيس أقصى ما لدى المتعلم من معلومات، حيث يترتب على نتائجها اجتياز هذا المتعلم للامتحانات ومن ثم ترقيته من صف دراسي لآخر، أو من مرحلة تعليمية إلى المرحلة التي تعلوها، أو تخرجه نهائياً من المرحلة التعليمية" (يوسف والرفاعي، 2005: 94).

كما ويعرف الاختبار المدرسي بأنه: "إجراء منظم لتحديد مقدار ما تعلمه الطالب من موضوع ما في ضوء الأهداف المحددة مسبقاً حيث يستفاد منه في تحسين أساليب التعلم وإجادة التخطيط وضبط التنفيذ وتقويم الإنجاز" (Gronlund, 1985).

ويعرفه بلوم (Bloom, 1976: 185) بأنه: "حكم لغرض ما على قيمة الأفكار والأعمال والحلول والمواد وغيرها، ويتضمن استخدام المحكات أو المستويات أو المعايير؛ لتقدير مدى كفاية الأشياء ودقتها وفعاليتها وقد يكون كمياً أو كيفياً".

ويعرفها الباحثان إجرائياً بأنها مقياس معياري وموضوعي يستخدم لتحديد مدى كسب المتعلم للمعلومات والمهارات من خلال إجابته على عينة من الأسئلة التي تم اختيارها لتمثل محتوى المادة الدراسة، ويمكن من خلال تطبيقها الوقوف عنج مواطن القوة والضعف، ومن ثم اتخاذ حكم حول مدى تحقي الأهداف التعليمية.

- **المرحلة المتوسطة:**

تعرف على أنها: "المدراس التابعة لوزارة التربية والتعليم السعودية، وتضم طلبة الصفوف الثلاث ما بعد المرحلة الأساسية، وتتراوح الفئة العمرية لهؤلاء الطلبة من 12 إلى 14 عاماً تقريباً" (العنزي، 2017: 9).

ويعرفها صنيدح (2016: 5) بأنها: "المرحلة الدراسية التي تشمل الصفوف الأربعة من مرحلة التعليم الأساسي وهي الصفوف السادس والسابع والثامن والتاسع حسب أنظمة وزارة التربية والتعليم الكويتية، وتتراوح أعمار الطلبة في هذه المرحلة بين (12- 15) سنة".

ويعرفها الباحثان إجرائياً بأنها إحدى مراحل التعليم العام بمنطقة تبوك بالمملكة العربية السعودية والتي تتضمن المراحل الثلاث الأول والثاني والثالث المتوسط.

- **تخطيط الاختبار:**

يعرفه الباحثان إجرائياً بأنه عبارة عن وضع خطة تفصيلية دقيقة تتمثل بدايتها في تحديد الغرض من الاختبار مروراً بإعداد جدول المواصفات وكتابة الفقرات للاختبار وانتهاءً بتصحيح الاختبار وتحليله.

- **إخراج الاختبار:**

يعرفه الباحثان إجرائياً بأنه عبارة عن المهام التي يقوم بها المعلم ليصل إلى ورقة الامتحان التي تقدم للطالب وتتضمن مجموعة من المهام منها: بناء وترتيب فقرات الاختبار والتأكد من وضوح الخط، وتناسقه، وطباعة الأسئلة وسلامة الصياغة ووضوح التعليمات.

- **تصحيح الاختبار:**

يعرفه الباحثان إجرائياً بأنه عبارة عن آلية التصحيح التي يتبعها المعلم في تصحيح الإجابات على الاختبار من إشارات أو علامات وغيرها من أساليب التغذية الراجعة والتعزيز.

2- الإطار النظري والدراسات السابقة:

الإطار النظري:

تؤكد النظريات التربوية الحديثة على جودة المخرجات المتوقعة للنظام التعليمي، ولا تكتفي بأهدافه وعملياته فقط؛ لأن المخرجات هي المؤشر الحقيقي على وصولنا للهدف المنشود في تحقيق مهمة التربية المتمثلة بإحداث تغيرات مرغوبة لدى المتعلمين سواء من الناحية المعرفية أو المهارية أو الانفعالية. وللتمكن من الحكم على مدى تحقيق النظام التعليمي لمخرجات عناصره الأساسية المتمثلة بالمنهاج والمعلم والطالب، لا بد من القيام بعملية

التقويم التربوي الذي يعتبر أحد المداخل لإصلاح العملية التعليمية التعلمية، الذي في ضوء نتائجه يتم إصدار الأحكام حول فاعلية البرامج والخطط لعمليات النظام التعليمي، فيمكن أن يكون التقويم وسيلة لتعميق الأهداف وتطويرها إذا استند إلى معايير علمية وموضوعية واضحة، كما يمكن أن يكون وسيلة للنهوض بالمنهاج بما يوفره من بيانات وأحكام للقيام بعملية تطويره على أسس واقعية، مما يزيد من فعالية تنفيذه. كما يمكن أن يكون وسيلة لتجديد طرق التدريس إذ أحسن المعلم الاستفادة من نتائجه ومؤشراته، ومن هنا لا بد من النظر إلى التقويم باهتمام خاص من حيث التخطيط له، وإعداد أدواته، واستخراج نتائجه وتحليله والاستفادة منها. (القاسم، 2016: 2).

يعد التقويم من أهم المداخل لتطوير العملية التعليمية، ومكوناً أساسياً من مكوناتها، فهو الوسيلة الأساسية التي يمكن بواسطتها التعرف إلى مدى النجاح في تحقيق الأهداف التربوية، وإلى كيفية الإفادة من جوانب القوة في العملية التعليمية التعلمية وتدعيمها وتعزيزها وذلك لتكرارها والحفاظ عليها، وإلى كيفية الإفادة من جوانب القصور لعلاجها ولتلاشيها، فنتائج التقويم يمكن اعتمادها في صنع القرارات ورسم العديد من السياسات التربوية. (الطراونة، 2012: 184).

أولاً- أهمية الاختبارات في التحصيلية:

مما لا شك فيه أن الاختبارات التحصيلية- إذا ما أحسن بناؤها واستخدامها- تكون عوناً لكل من الطالب والمعلم وصانع القرار على السواء.

أ- أهمية الاختبارات التحصيلية بالنسبة للمعلم:

تتمثل أهمية الاختبارات التحصيلية بالنسبة للمعلم في أنها تساعد في التعرف على المستوى التحصيل الدراسي الذي وصل إليه التلاميذ وبالتالي مراقبة تقدم العملية التعليمية من خلال ما يحدث لهم من تحسن أو تأخر في التحصيل الدراسي.

ب- أهمية الاختبارات بالنسبة لصانع القرار:

لا شك أن نتائج الاختبارات التحصيلية تزود صانع القرار التربوي بمعلومات جيدة يستخدمها في إصدار العديد من القرارات الإدارية مثل الترفيع أو النقل من فرقة دراسية إلى فرقة أعلى منها (سليمان وعلام، 2010: 45).

ج- أهمية الاختبارات التحصيلية بالنسبة للطالب:

للاختبارات تأثيرات مفيدة في تعلم التلاميذ ومن هذه التأثيرات ما يلي:

1- تزيد الاختبارات دافعية التلاميذ لتعلم المواد الدراسية خصوصاً في حالة الامتحانات الرسمية.
2- تشجع الاختبارات على مراجعة ومذاكرة المعلومات التي سبق تعلمها.
3- تؤثر الاختبارات في العمليات المعرفية التي يستخدمها التلاميذ حين يتعلمون.
4- توفر الاختبارات خبرات تعلم في ذاتها وبذاتها.
5- توفر الاختبارات تغذية راجعة عن مدى تعلم التلاميذ (جابر، 2002: 45).

ثانياً- تصنيف الاختبارات التحصيلية:

هناك أكثر من طريقة لتصنيف الاختبارات التحصيلية ومنها:

أ- التصنيف وفق طريقة أداء الطلبة: اختبار شفوي، اختبار تحريري، اختبار أدائي.
ب- التصنيف وفق موعد التطبيق: اختبارات أولية، اختبارات بنائية، اختبارات نهائية.
ج- التصنيف حسب الطريقة التي تُفسر بها الدرجات: اختبار مرجعي المحك، واختبار مرجعي المعيار.
د- التصنيف وفق طريقة التصحيح: اختبارات موضوعية، واختبارات غير موضوعية.

هـ- التصنيف وفق استخدام الكتاب: اختبارات الكتاب المفتوح، اختبارات الكتاب المغلق.

و- التصنيف وفق محتوى الاجابة: لفظي وغير لفظي (مراد وسليمان، 2005: 80).

ثالثاً- شروط الاختبار الجيد:

هناك مجموعة من الشروط يجب أن تتوافر في الاختبار حتى يكون اختباراً جيداً مؤدياً للغرض الذي وضع من أجله. وهذا الاختبار لا يكتمل إلا إذا توافرت معلومات عن صلاحيته باعتباره أداة للقياس تشتمل على شروط معينة أهمها: الصدق والثبات والموضوعية، وفيما يلي توضيح لهذه الشروط:

1- الصدق:

إننا نعني بصدق الاختبار مدى قياس الاختبار- فعلاً- للاستعداد أو الخاصية أو السمة التي وضع ليقيسها. فإذا وضعنا اختباراً بقصد قياس الذكاء فإلى أي حد يقيس هذا الاختبار- بالفعل- خاصية الذكاء وليس شيئاً غيرها؟ ولهذا، إن صدق الاختبار يعتبر صفة أساسية ينبغي توافرها في الاختبار؛ وإلا فقد الاختبار قيمته كوسيلة لقياس الاستعداد أو الخاصية المقصود به قياسها (الشربيني وآخرون، 2010: 389).

2- الثبات:

يقصد بمفهوم ثبات الدرجات مدى خلوها من الأخطاء غير المنتظمة التي تشوب القياس، أي مدى قياس الاختبار للمقدار الحقيقي للسمة التي يهدف لقياسها، فدرجات الاختبار تكون ثابتة إذا كان الاختبار يقيس سمة معينة قياساً متسقاً في الظروف المتباينة التي قد تؤدي إلى أخطاء القياس (علام، 2011: 131).

3- الموضوعية أو دقة القياس:

ويُحكم على الاختبار بالموضوعية إذا كانت أسئلته وإجاباته محددة، بحيث يكون للسؤال الواحد جواب واحد لا يختلف عليه المصححون، هذا وترتبط موضوعية الاختبار بمفهوم العدل في إعطاء الدرجات، مما يزيد من ثقة المفحوصين في نتائجهم.

شروط ثانوية للاختبار الجيد:

هناك بعض الشروط الثانوية التي ينبغي مراعاتها عند إعداد الاختبارات، وهذه الشروط تشمل:

1- **سهولة التطبيق:** حيث أنه كلما كان إجراء الاختبار سهلاً كلما كان ذلك أفضل، إذ إن صعوبة التطبيق قد يكون عائقاً يمنع تحقيق الموضوعية والثبات والصدق.

2- **سهولة تقدير الدرجة:** وهو يعني سهولة تصحيح الاختبار وعادة ما يكون الاختبار سهلاً في تصحيحه إذا كان سهلاً في تطبيقه. وترتبط سهولة التصحيح بموضوعية الاختبار.

3- **سهولة التفسير:** ترتبط سهولة التفسير بمعاملين وهما: معايير الاختبار وتعليماته.

4- **الاقتصاد أو التكلفة:** ينبغي ألا يهمل المعلم الجانب الاقتصادي عند اختياره للاختبارات التي يستخدمها، فكلما كانت تكلفة إعداد الاختبار قليلة يكون ذلك أفضل (فتح الله، 2016: 298).

ثانياً- الدراسات السابقة:

- دراسة المالكي (2017). هدفت إلى التعرف على واقع الاختبارات الوطنية في ضوء معايير جودة التقويم، ومدى تحقيقها للمعايير الخاصة بالمناهج القطرية في المواد الدراسية الأربع (اللغة العربية، اللغة الإنجليزية،

الرياضيات، والعلوم)، في الصفوف من الرابع وحتى السادس بالمدارس الابتدائية المستقلة. استخدم الباحث المنهج الوصفي التحليلي. وتكون مجتمع الدراسة من (12) ورقة اختبارية من الاختبارات الوطنية، وفي المواد الدراسية الأربع وللصفوف الثلاثة المتقدمة، التي طبقت خلال السنتين (2011- 2012/2012- 2013). أشارت نتائج الدراسة إلى وجود محور غير متحقق في مادتي العلوم والرياضيات وهو محور "تعليمات الاختبار"؛ حيث كان هناك قصور في بند "زمن الاختبار الكلي"، و"الدرجة المعطاة لكل سؤال"، و"تناسب عدد الأسئلة مع زمن الاختبار الكلي في المواد الرئيسة الأربعة" ولجميع الصفوف الدراسية.

- دراسة المسعودين (2017). هدفت إلى التعرف على درجة التزام معلمي المرحلة الأساسية في الأردن بالمعايير الوطنية في بناء اختبارات التحصيل من وجهة نظر المديرين والمشرفين التربويين. ولتحقيق أهداف الدراسة، أعد الباحث استبانة بالمعايير الوطنية في بناء اختبارات التحصيل، وتكونت عينة الدراسة من (19) مشرفاً ومشرفة و (36) مديراً و (42) مديرة. لقد أظهرت نتائج الدراسة أن درجة التزام معلمي المرحلة الأساسية في الأردن بالمعايير الوطنية جاءت بدرجة متوسطة، وهناك فروق ذات دلالة إحصائية بين معدلات التزام معلمي المرحلة الأساسية في الأردن بالمعايير الوطنية تُعزى لمتغيرات (الوظيفة، والجنس، والخبرة).
- دراسة الرويلي (2016). وهدفت إلى تحليل مهارات بناء الاختبارات التحصيلية لدى الطلاب المعلمين من وجهة نظرهم والمختصين في المملكة العربية السعودية، وذلك باستخدام قائمة لأهم مهارات القياس والتقويم واستخدام هذه القائمة في الكشف عن آراء الطلاب المعلمين حول درجة إتقانهم لهذه المهارات. وقد تم أيضاً استخدام اختبار من نوع (صواب وخطأ) لقياس معرفتهم في بناء الاختبارات التحصيلية. لقد تم تطبيق الأداة على (470) طالباً تم اختيارهم بطريقة عشوائية و(13) استاذاً ممن يُدرِّسون المادة في كليات المعلمين و(24) استاذاً من مختصي القياس والتقويم في الجامعات العربية، وكان أبرز نتائج الدراسة ما يلي: تدني معرفة الطلاب بمهارات القياس والتقويم المتعلقة ببناء الاختبارات وأن درجة التركيز على المهارات أثناء تدريس مادة التقويم التربوي كانت متوسطة.
- دراسة النفيعي (2016). هدفت إلى معرفة مدى إتقان أعضاء مشروع الاختبارات التحصيلية في كل منطقة من المناطق المنفذ بها المشروع لمهارات بناء الاختبارات التحصيلية ومعرفة جوانب القوة والضعف بها في جمهورية مصر العربية. (تم استخدام المنهج الوصفي وتكونت العينة من (146) عضو من أعضاء المشروع تم استخدام اختبار لتقدير مهارات بناء الاختبارات التحصيلية وكانت أهم النتائج تدني مستوى إتقان أعضاء المشروع للمهارات الأساسية لبناء الاختبارات التحصيلية من حيث (التخطيط –أعداد الفقرات –التصحيح).
- دراسة عرفات (2015). هدفت إلى معرفة مدى امتلاك أساتذة التعليم الثانوي في المدارس الثانوية في ثانويات الوادي بولاية الواد في الجزائر لكفايات إعداد الاختبارات التحصيلية الجيدة. تم استخدام المنهج الوصفي بأسلوبه الاستكشافي ولتحقيق الأهداف أعدت أداة للدراسة طبقت على عينة بلغت (٢٤١) استاذ ثانوي لمعرفة مدى امتلاكهم لكفايات (التخطيط- الإعداد والتقنين- التحليل والتطبيق) في الاختبارات التحصيلية وتوصلت الدراسة إلى أن معظم الأساتذة لا يمتلكون كفايات إعداد الاختبارات التحصيلية.
- دراسة الكلثم (2015). هدفت الدراسة إلى تحديد معايير جودة ورقة الاختبار، ثم التعرف على مدى توافر هذه المعايير في ورقة الاختبار في اختبارات تخصص تربية إسلامية في جامعتي أم القرى والدمام. استخدم الباحث المنهج الوصفي، وقد قام الباحث بتحليل (30) ورقة اختبار بواقع (15) لكل جامعة في ضوء مواصفات ورقة الاختبار من حيث الشكل، ومواصفات ورقة الاختبار من حيث المضمون، وجودة الأسئلة الاختبار مقالية وموضوعية، وقد توصل الباحث إلى النتائج التالية: 1- تم تحديد (58) معيار جودة لورقة

الاختبار. 2- أن نسبة توافر مواصفات ورقة الاختبار تراوحت ما بين (93% و75%) في أوراق الاختبار الخاصة بكلية الشريعة والدعوة بجامعة أم القرى وهي نسبة عالية إلى حد ما. أما بالنسبة لأوراق اختبارات كلية التربية، جامعة الدمام فقد تراوحت نسبة التوافر ما بين (93%، و51%)، وهي نسبة متوسطة. 3- أن مستوى التذكر قد حظي بنسبة 87%، وذلك في أسئلة أوراق الاختبار في كل من كلية الشريعة والدعوة جامعة أم القرى، وكلية التربية جامعة الدمام. واحتلت بذلك المرتبة الأولى في أسئلة المستويات المعرفية.

- دراسة العمري (2015). هدفت إلى الكشف عن مدى المعرفة التي تتوافر لدى معلمي العلوم للمرحلة الأساسية في لواء المزار الشمالي بمدينة إربد في الأردن بمبادئ بناء الاختبارات التحصيلية، تكونت عينة الدراسة من (103) معلماً ومعلمة، ولتحقيق أهداف الدراسة طور الباحث استبانة لقياس مدى معرفة المعلمين بمبادئ بناء الاختبارات التحصيلية. أظهرت النتائج أن معرفة معلمي العلوم بمبادئ بناء الاختبارات التحصيلية جاءت بمستوى متوسط، وأن أكثر مجالات بناء الاختبارات التحصيلية استخداماً هو مجال تطبيق الاختبار، كما أظهرت النتائج وجود فروق ذات دلاله إحصائية في مدى المعرفة التي تتوافر لدى المعلمين بمبادئ بناء الاختبارات التحصيلية تُعزى لمتغير الخبرة والجنس.
- دراسة عتوم (2014). هدفت إلى التعرف على مدى مطابقة الاختبارات المدرسية التحصيلية من إعداد معلمي ومعلمات مدارس مديرية التربية والتعليم في محافظة جرش مع معايير الاختبار الجيد. تم تطبيق الدراسة على عينة مؤلفة من (116)، اختباراً تحصيلياً في نهاية الفصل الدراسي الثاني من العام الدراسي (2012/ 2013) في مباحث العلوم والفيزياء والكيمياء والأحياء وعلوم الأرض لقد اعتمدت الدراسة على الاستبانة في جمع البيانات. وتوصلت الدراسة إلى أن الاختبارات التحصيلية من إعداد المعلمين تتفق مع معايير الاختبار الجيد بدرجة متوسطة، وأن هناك فروقاً ذات دلالة إحصائية بين المعلمين والمعلمات في مدى تطابق الاختبار مع معايير الاختبار الجيد لصالح المعلمات.
- دراسة ساعد (2013). وهدفت إلى التعرف على مدى توافر معايير الاختبار التحصيلي الجيد في الاختبارات التي يعدها أساتذة التعليم الجامعي في ولاية الجزائر من حيث كتابة تعليمات الاختبار وإعداد فقراته وكذا المعايير اللازمة في مجال طباعة الاختبار وإخراجه. تم الاعتماد على المنهج الوصفي التحليلي، حيث تم القيام بتحليل محتوى (36) اختباراً تحصيلياً للسنة الجامعية 2013/2012، وتوصلت نتائج الدراسة إلى أن معايير كتابة تعليمات الاختبار التحصيلي الجيد وإعداد فقراته جاءت بنسب ضعيفة، وأن المعايير المتعلقة بطباعة الاختبار وإخراجه كانت بنسبة متوسطة. كما توصلت نتائج الدراسة أيضاً إلى أن الأساتذة أكثر استخداما للأسئلة المقالية مقارنة مع الأسئلة الموضوعية.
- دراسة قدار (2013). هدفت إلى تحديد أهم معايير الاختبارات التحصيلية الجيدة، ومدى توافرها في الاختبارات التحصيلية الوزارية للمواد الاجتماعية للأعوام 2006- 2010 في الجمهورية السمنية. لقد تم استخدام استبانة اشتملت على خمسة معايير، وتكونت عينة البحث من أسئلة الاختبارات التحصيلية اشتملت على (120) سؤالاً بمعدل (6) أسئلة لكل ورقة اختبارية، وكان مجمل الأسئلة الفرعية (253). واستخدمت النسبة المئوية ومعادلة كوبر كوسائل إحصائية في البحث. وكانت نتائج الدراسة على النحو الآتي: أظهر التحليل في ضوء مستويات المجال المعرفي لبلوم أن نسبة التذكر (55.32%) والفهم (35.67%) والتطبيق (6%) والتحليل (1.6%).

التعليق على الدراسات السابقة:

تم عرض بعض الدراسات التي اهتمت بتقويم الاختبارات التحصيلية التي يعدها المعلمون وسيتم التعليق عليها بالنقاط التالية:

- هناك وفرة في الدراسات التي اهتمت بتقويم امتحانات المراحل الجامعية والثانوية يقابلها ندرة في الدراسات التي تهتم بتقويم اختبارات المرحلة المتوسطة.
- اهتمت معظم الدراسات بتقويم الامتحانات في ضوء النواحي الفنية والإخراجية للاختبار بالإضافة إلى مدى شموله للمستويات المعرفية تبعا لتصنيف بلوم، ومدى تنوع الأسئلة.
- المنهج المستخدم في الدراسات السابقة هو المنهج الوصفي وهو ما اتبعته الباحثان.
- اتفقت أغلب الدراسات على أن مستوى التذكر هو المستوى الذي حصل على نسبة عالية من الأسئلة يليه مستوى الفهم بنسبة أقل، وحصل مستوى التطبيق على نسب متدنية مقارنه بالمستويات السابقة.

وتأتي الدراسة الحالية امتداداً للدراسات السابقة، وتتميز بأنها تستهدف المرحلة المتوسطة وهي مرحلة تعاني من ندرة الدراسات في مجال تقويم الامتحانات- على حد علم الباحثين- وأيضاً تتميز بأنها تناقش ثلاثة محاور مهمة في بناء الاختبارات وهم محور التخطيط، ومحور الإخراج، ومحور التصحيح.

3- منهجية البحث وإجراءاته:

منهج البحث:

اعتمد الباحثان على المنهج الوصفي التحليلي والذي يستهدف الوصف الكمي والمنظم للمحتوى وبذلك يصل الباحث إلى إصدار حكم معين على الاتجاه الغالب في قضية معينة (العساف، 2010: 217).

مجتمع البحث وعينته:

¬ مجتمع البحث:

جميع معلمات العلوم للمرحلة المتوسطة في منطقة تبوك والبالغ عددهن (138) معلمة موزعات على (58) مدرسة في منطقة تبوك.

¬ عينة البحث:

تم اختيار (30) ورقة بطريقة عشوائية من أوراق اختبار النصف الأول لمادة العلوم صف أول متوسط لغرض فحص مدى التزام المعلمات بمعايير إخراج الاختبارات المدرسية، وتم أيضاً اختيار (100) معلمة بطريقة عشوائية لغرض التعرف على مدى التزام معلمات العلوم في المرحلة المتوسطة بمعايير التخطيط للاختبارات وتصحيحها.

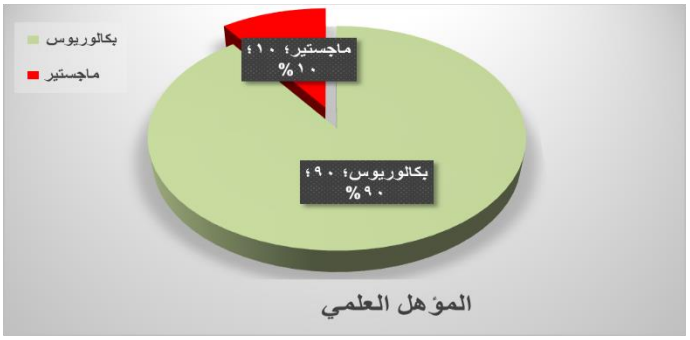
والجداول التالية تبين خصائص وسمات عينة الدراسة كما يلي:

1- المؤهل العلمي:

يبين جدول رقم (1) أن (90.0%) من عينة الدراسة مؤهلهم العلمي "بكالوريوس"، و(10.0%) مؤهلهم العلمي "ماجستير".

جدول رقم (1) توزيع عينة الدراسة حسب متغير المؤهل العلمي

المؤهل العلمي	التكرار	النسبة المئوية
بكالوريوس	90	%90
ماجستير	10	%10
المجموع	100	%100



رسم بياني يوضح نسبة الماجستير إلى البكالوريوس في المؤهل العلمي رقم (1)

2- سنوات الخبرة:

يبين جدول رقم (2) أن (52.0%) من عينة الدراسة تراوحت سنوات الخبرة لهم 5 سنوات فأقل من الخبرة، و(28.0%) يتملكون من 6 إلى 10 سنوات من الخبرة، و (20.0%) لديهم أكثر من 10 سنوات من الخبرة.

جدول رقم (2) توزيع عينة الدراسة حسب متغير سنوات الخبرة

سنوات الخبرة	التكرار	النسبة المئوية
5 سنوات فأقل	52	%52
6- 10 سنوات	28	%28
أكثر من 10 سنوات	20	%20
المجموع	100	%100

أدوات البحث:

تمثلت أداتا البحث في:

1- استبانة موجهة للمعلمات لمعرفة مدى التزامهن بمهارة تخطيط وتصحيح للاختبارات التحصيلية.
2- استمارة معايير جودة إخراج الاختبار، واستخدم تحليل المحتوى كأداة لجمع المعلومات.

أولاً: الأداة المتعلقة بممارسة معلمات العلوم في المرحلة المتوسطة لمعايير التخطيط والتصحيح للاختبار.

ومر بناء هذه الأداة بالخطوات التالية:

1. تم الاطلاع على الدراسات والبحوث في مجال بناء وتقويم الاختبارات التحصيلية مثل دراسة (العمري، 2015)، (النفيعي، 2016) وتم الاستفادة منها. كما تم الاطلاع على توصيات بعض الندوات والمؤتمرات، وبعض المراجع والمؤلفات في مجال القياس والتقويم.
2. وُضِعت الأداة في صورتها الأولية حيث احتوت على محورين هما: محور التخطيط ومحور التصحيح.

صدق الأداة (الاستبانة):

أولاً/ الصدق الظاهري للأداة (صدق المحكمين):

تم عرض الأداة على مجموعة من المحكمين من ذوي الخبرة في مجال القياس والتقويم التربوي في الجامعات في المملكة العربية السعودية والجامعات العربية، وعددهم (6) محكمين، وقد طُلِب من المُحكمين إبداء آرائهم في مدى ملاءمة العبارات لقياس ما وضعت لأجله، ومدى وضوح صياغة العبارات ومدى مناسبة كل عبارة للمحور الذي تنتمي إليه. ومدى كفاية العبارات لتغطية كل محور من محاور متغيرات الدراسة الأساسية هذا بالإضافة إلى اقتراح ما يرونه ضروريا من تعديل صياغة العبارات أو حذفها، وإضافة المحاور التي يرون ضرورة إضافتها للأداة،

وقد أقر المحكمون تلك المحاور التي وضِعت، وكانت هناك ملاحظات على بعض الفقرات التي تتضمنها المحاور، ففي محور التخطيط أوصوا بحذف الفقرات (9، 10، 11، 13) أما في محور التصحيح فقد تمت التوصية بحذف الفقرات (17، 18، 20، 21) واستبدالهم بعبارة الابتعاد عن الذاتية قدر الامكان في التصحيح. أما الفقرات (22، 23) فتم حذفها لأنها لا تنتمي للمجال موضع الاهتمام.

وبعد أن تمت التعديلات في ضوء ملاحظات المحكمين على الأداة خرجت الأداة بصورتها النهائية.

وفيما يلي وصفًا مفصلاً لمحتويات الاستبانة في صورتها النهائية:

أ- **محور التخطيط**: ويتضمن (10) فقرات تغطي مهارة التخطيط والتهيئة للاختبارات التحصيلية.

ب- **محور التصحيح**: ويتضمن (11) فقرة تغطي مهارات تصحيح الاختبارات التحصيلية.

ثانياً/ صدق الاتساق الداخلي لفقرات الاستبانة:

تم حساب معاملات الاتساق الداخلي لفقرات الاستبيان على عينة استطلاعية البالغ حجمها (30) مفردة، وذلك بحساب معاملات الارتباط بين كل فقرة والدرجة الكلية للمحور التابعة له، ويبين جدول رقم (3) معاملات الارتباط بين الدرجة على كل فقرة من فقرات المحور والدرجة الكلية على الأداة.

جدول (3) الصدق الداخلي لفقرات الاستبانة

المحور الأول: التخطيط			المحور الثاني: التصحيح		
رقم الفقرة	معامل الارتباط	مستوى الدلالة	رقم الفقرة	معامل الارتباط	مستوى الدلالة
1	0.706	0.000	1	0.712	0.000
2	0.554	0.000	2	0.765	0.000
3	0.829	0.000	3	0.568	0.001
4	0.653	0.000	4	0.654	0.000
5	0.671	0.000	5	0.746	0.000
6	0.730	0.000	6	0.532	0.002
7	0.593	0.001	7	0.624	0.000
8	0.596	0.001	8	0.735	0.000
9	0.624	0.000	9	0.433	0.017
10	0.781	0.000	10	0.735	0.000
			11	0.673	0.000

قيمة r الجدولية عند مستوى دلالة 0.05 ودرجة حرية "28" تساوي 0.361

ثالثاً/ ثبات الأداة:

لمعرفة ثبات أداة البحث فقد تم تطبيقها على عينة استطلاعية عددها (30) معلمة من معلمات المرحلة المتوسطة في تبوك، وتم حساب معامل الثبات بطريقتين هما طريقة التجزئة النصفية ومعامل ألفا كرونباخ، كما في الجدول رقم (4)

جدول رقم (4) معامل الثبات لأداة البحث

المحور	عنوان المحور	العدد	معامل الثبات بالتجزئة النصفية	معامل الثبات بطريقة ألفا كرونباخ
الأول	التخطيط	10	0.837	0.837
الثاني	التصحيح	11	0.898	0.881
الأداة ككل		21	0.858	0.867

الأداة الثانية: قائمة معايير بناء وإخراج الاختبار:

بعد الاطلاع على الدراسات والبحوث في مجال بناء وتقويم الاختبارات التحصيلية مثل دراسة (عتوم، 2014) التي تم الاستفادة منها، وتوصيات بعض الندوات والمؤتمرات، وبعض المراجع والمؤلفات في مجال القياس والتقويم. تم إعداد الأداتين الأولى تتعلق بإخراج الاختبار في صورتها الأولية حيث احتوت على محورين هما معايير بناء وإخراج الاختبار.

أولاً/ صدق الأداة:

تم عرض القائمة بمعايير بناء وإخراج الاختبار على مجموعة من المحكمين عددهم (6) للتأكد من صدقها وانتماء الفقرات للمجال المراد قياسه وقد أوصوا بزيادة فقرات المزاوجة وهو ما تم تنفيذه. ولم تكن هناك مقترحات أخرى بشأنها، وبعد أن تمت التعديلات في ضوء ملاحظات المحكمين على الأداتين خرجت الأداتان بصورتهما النهائية وكان عدد فقراتها (29) فقرة (محور إخراج الاختبار (9) فقرات، ومحور بناء الاختبار (20) فقرة.

ثانياً/ ثبات الأداة:

الثبات بطريقة إعادة الاختبار:

تم إعادة تحليل عينة من الاوراق بواسطة باحثة أخرى وتبين أن معامل الارتباط بين مرتي التطبيق كان (0.90)

جدول رقم (5) الثبات بطريقة إعادة الاختبار

المحور	عنوان المحور	معامل الثبات
الأول	مدى استيفاء الاختبارات المحللة للمعايير الفنية لإخراج الاختبار	0.93
الثاني	مدى استيفاء الاختبارات المحللة للمعايير بناء الفقرات لإخراج الاختبار	0.94
للأداة ككل		0.90

رابعاً: خطوات (إجراءات البحث) والأساليب الإحصائية المستخدمة:

خطوات (إجراءات البحث):

أ- تحليل الاختبارات:

1. جُمِعت اختبارات مادة العلوم للصف أول متوسط الفصل الدراسي الأول من (30) مدرسة من مدارس منطقة تبوك، بعد التحقق من صدق الأداة وثباتها تم البدء بإجراء عملية التحليل.
2. بعد انتهاء عملية تحليل الاختبارات، تم إدخال بيانات التحليل بعد ترميزها؛ حتى يسهل إدخالها إلى الحاسوب باستخدام برنامج الحزم الإحصائية للعلوم الاجتماعية (SPSS)، ومن ثم إجراء التحليلات الإحصائية.
3. تحليل النتائج، وتفسيرها، وبيان دلالتها التربوية.

ب- الاستبانة:

1. تم تصميم استبانة تحوي محورين محور التخطيط للاختبار، ومحور تصحيح الاختبار وفق معايير وشروط الاختبار الجيد.
2. بعد التحقق من صدق الأداة وثباتها واعتمادها من المشرف على الرسالة تم توزيع الاستبانة على العينة وتجميع الاستجابات.
3. ثم تم إدخال استجابات المعلمات بعد ترميزها؛ حتى يسهل إدخالها إلى الحاسوب باستخدام برنامج الحزم الإحصائية للعلوم الاجتماعي (SPSS)، ومن ثم إجراء التحليلات الإحصائية.
4. تحليل النتائج، وتفسيرها، وبيان دلالتها التربوية.

خامساً: المعالجات الإحصائية:

وفيما يلي مجموعة من الأساليب الإحصائية المستخدمة في تحليل البيانات:

1. تم حساب التكرارات والنسب المئوية.
2. المتوسط الحسابي Mean وذلك لمعرفة مدى ارتفاع أو انخفاض استجابات أفراد الدراسة عن كل عبارة من عبارات متغيرات الدراسة الأساسية.
3. تم استخدام الانحراف المعياي (Standard Deviation) للتعرف على مدى انحراف استجابات أفراد الدراسة لكل عبارة من عبارات متغيرات الدراسة ولكل محور من المحاور الرئيسية عن متوسطها الحسابي.
4. معامل ارتباط بيرسون لقياس صدق الفقرات والعلاقات بين أبعاد الدراسة.
5. معادلة سبيرمان براون للثبات.
6. معامل الثبات (ألفا كرونباخ).
7. معامل الثبات (التجزئة النصفية).
8. معامل الثبات (إعادة الاختبار).
9. اختبار تحليل التباين الثنائي لحساب الفروق بين متوسطات ثلاث عينات مستقلة فأكثر.

4- عرض النتائج ومناقشتها:

أولاً: النتائج الخاصة بالسؤال الأول: ما درجة التزام معلمات العلوم للمرحلة المتوسطة بمهارات التخطيط للاختبارات وفق معايير تخطيط للاختبار؟

للإجابة على هذا السؤال تم استخدام اختبار t للعينة الواحدة والنتائج مبينة في جدول رقم (6) والذي يبين آراء أفراد عينة الدراسة في فقرات المحور الأول (التخطيط) مرتبة من الأعلى إلى الأسفل حسب المتوسط الحسابي هي كما يلي:

جدول (6) المتوسطات الحسابية والانحرافات المعيارية لإجابات العينة على درجة التزام معلمات العلوم للمرحلة المتوسطة بمعايير التخطيط للاختبار (مرتبة تنازلياً).

م	الفقرات	المتوسط الحسابي	الانحراف المعياري	درجة الموافقة	الترتيب
7	أخبر الطلاب بموعد الاختبار قبل الاختبار بوقت كاف.	4.73	0.51	عالية جداً	1
6	أحرص على أن يكون زمن الاختبار مناسباً لطول الأسئلة.	4.61	0.57	عالية جداً	2
4	أحرص على تنوع فقرات الاختبار (موضوعية –مقالية).	4.52	0.72	عالية جداً	3
3	أحرص على أن تغطي فقرات الاختبار أهداف المقرر.	4.48	0.68	عالية جداً	4
5	أحدد مستوى المجال الذي يقيسه الاختبار في ضوء قدرات الطلبة وطبيعة المادة التعليمية.	4.43	0.68	عالية جداً	5
9	استخدم الأساليب التربوية التي تخفف من قلق الاختبار لدى الطلبة.	4.34	0.78	عالية جداً	6
10	أحدد طريقة وآلية تصحيح الاختبار لضبط أثر التخمين وأثر الهالة.	4.28	0.81	عالية جداً	7
2	أحدد نوع الاختبار بما يتناسب مع الموقف التعليمي.	4.26	0.81	عالية جداً	8
8	أحدد الأهداف العامة والخاصة من الاختبار التحصيلي. قبل إعداده.	4.14	0.98	عالية	9
1	أقوم بتصميم جدول مواصفات للاختبار.	3.38	1.09	متوسطة	10
	الكلي	4.32	0.47	عالية جداً	

يظهر من الجدول (6) أن المتوسطات الحسابية لفقرات مجال تخطيط الاختبار تراوحت بين عالية إلى متوسطة بمتوسط حسابي (4.73- 3.38) مما يدل على أن التزام معلمات المرحلة المتوسطة لمعايير التخطيط للاختبارات التحصيلية وفق معايير وشروط الاختبار الجيد عالية. ويعلل الباحثان الدرجة العالية من الممارسة بسبب تلقي المعلمات العديد من الدورات الخاصة ببناء الاختبارات التحصيلية، وإقامة ورش تدريبية للمعلمات في هذا المجال، بالإضافة إلى أن كثير من المعلمات قد تعرضن لدراسة مساقات تربوية خاصة بالقياس والتقويم التربوي وتتفق نتائج هذه الدراسة مع نتائج دراسة (الزهراني، 2011) التي دلت على ارتفاع مهارات التخطيط لدى المعلمين والمشرفين. واختلفت النتائج مع نتائج دراسة (العمري، 2015) ودراسة (الرويلي، 2016) فقد دلت نتائجها على أن إلمام المعلمين بمهارات التخطيط والتهيئة للاختبارات جاءت متوسطة، كما اختلفت مع نتيجة دراسة (عرفات، 2015) التي أشارت نتائجها إلى ضعف امتلاك المعلمين لكفايات (التخطيط- الإعداد والتقنين- التحليل والتطبيق) في الاختبارات التحصيلية ودراسة الكثلم (2015) التي دلت على ضعف بمستوى التخطيط والتهيئة إذ بلغت نسبة إتقان هذا المجال (36.99%)

ثانياً: النتائج الخاصة بالسؤال الثاني: ما درجة التزام معلمات العلوم للمرحلة المتوسطة بمعايير تصحيح الاختبار؟

للإجابة على هذا السؤال تم استخدام اختبار t للعينة الواحدة والنتائج مبينة في جدول رقم (7) والذي يبين آراء أفراد عينة الدراسة في فقرات المحور الثاني (التصحيح) مرتبة من الأعلى إلى الاسفل حسب الوزن النسبي هي كما يلي:

جدول (7) المتوسطات الحسابية والانحرافات المعيارية لإجابات العينة على درجة التزام معلمات العلوم للمرحلة المتوسطة بمعايير تصحيح الاختبار(مرتبة تنازلياً).

م	الفقرات	المتوسط الحسابي	الانحراف المعياري	درجة الموافقة	الترتيب
1	التصحيح دائماً بقلم مغاير للون كتابة الطلاب.	4.78	0.44	عالية جداً	1
4	الابتعاد عن الذاتية قدر الإمكان في التصحيح.	4.57	0.65	عالية جداً	2
6	أكتب الدرجة المستحقة لكل سؤال.	4.55	0.76	عالية جداً	3
3	التصحيح من خلال نماذج الإجابة المناسبة للاختبار.	4.52	0.67	عالية جداً	4
7	أقبل أي جواب صحيح آخر يقدمه الطالب غير ما ورد في الكتاب أو الحصة الدراسية.	4.47	0.81	عالية جداً	5
2	أتجنب التصحيح وأنا مرهق أو أمر بضغوط نفسية.	4.42	0.86	عالية جداً	6
5	أصحح الأسئلة المقالية بالطريقة الشمولية.	4.31	0.82	عالية جداً	7
9	أجيب عن أسئلة الطلبة بعد الانتهاء من تطبيق الاختبار.	4.19	0.98	عالية	8
11	ألتزم بفترة محددة لتصحيح أوراق الاختبار.	3.84	1.00	عالية	9
10	أسلم أوراق الاختبار للطلبة بعد التصحيح.	3.79	1.14	عالية	10
8	أكتب ملاحظاتي على إجابات الطلبة عند تصحيح الأسئلة المقالية.	3.51	1.07	عالية	11
	جميع الفقرات	4.27	0.45	عالية جداً	

يظهر من الجدول (7) أن مجال تصحيح الاختبار قد حصل على متوسط (4.27) أي بدرجة (عالية جداً)، وبخصوص الفقرات فقد تراوحت المتوسطات الحسابية لفقرات مجال تصحيح الاختبار بين (4، 78- 3، 51) وبصفة عامة يتبين أن المتوسط الحسابي لجميع فقرات المحور الثاني (التصحيح) تساوي 4.27 مما يدل على أن التزام معلمات المرحلة المتوسطة بمعايير التصحيح للاختبارات التحصيلية وفق معايير وشروط الاختبار الجيد عالية.

وقد يكون هذا عائد لارتباط مهارات التصحيح عند المعلمات بالأمانة العلمية والعدالة، وأيضاً لاهتمام الإدارة المدرسية بعملية التصحيح إذ أنه غالباً ما يقوم بتصحيح الاختبارات لجنة مكونة من مصححة ومراجعة ومدققة لضمان عدم وجود أخطاء في عملية التصحيح.

ونلاحظ في هذا المحور أيضاً ارتفاع متوسط استجابات محور التصحيح وتتفق مع ذلك دراسة (العمري، 2015) التي دلت على ارتفاع محور التصحيح وتخالف دراسة (النفيعي، 2016) التي دلت على تدني مهارات التصحيح لدى أعضاء مشروع الاختبارات التحصيلية من المعلمين والمشرفين.

ثالثاً: النتائج الخاصة بالسؤال الثالث: ما مدى التزام معلمات المرحلة المتوسطة بمعايير إخراج الاختبار؟

ويتفرع من هذا السؤال الأسئلة الفرعية التالية:

أ- ما مدى استيفاء اختبارات معلمات المرحلة المتوسطة للمعايير الفنية لإخراج الاختبار؟

للإجابة على هذا السؤال تم استخدام النسب المئوية والنتائج مبينة في جدول رقم (8) والذي يبين نتائج تحليل عينة الدراسة في فقرات المحور الأول: مدى استيفاء الاختبارات المحللة للمعايير الفنية لإخراج الاختبار مرتبة من الأعلى إلى الأسفل حسب النسب المئوية هي كما يلي:

جدول رقم (8) مدى استيفاء الاختبارات المحللة للمعايير الفنية لإخراج الاختبار

م	الفقرات	مستوفي%	مستوفي إلى حد ما %	غير مستوفي%	الترتيب
1	يوجد على ورقة الاختبار التعليمات العامة. اليوم التاريخ. الصف. المادة. المدرسة.	%96.7	%3.3	%0.0	1
5	الطباعة واضحة ومقروءة.	%96.7	%3.3	%0.0	1
9	كتابة الفقرة كاملة في نفس الصفحة.	%96.7	%0.0	%3.3	1
6	إعطاء الفقرات أرقاماً متسلسلة.	%93.3	%6.7	%0.0	3
8	أضع فراغ مناسب للإجابة.	%93.3	%6.7	%0.0	3
7	وضوح الخط وتناسقه ومناسبة المسافة بين الأسطر والكلمات.	%73.3	%20.0	%6.7	4
2	تخصيص الصفحة الأولى من كراس الاختبار لكتابة اسم الطالب ودرجاته وبقية البيانات الأساسية.	%70	%0.0	%30	5
3	خلو الاختبار من الأخطاء اللغوية والمطبعية.	%66.7	%30.0	%3.3	6
4	على ورقة الاختبار يوجد علامة كل سؤال.	%53.3	%6.7	%40	7
	المتوسط العام للمجال.	%.82	%85.2	%83.3	

في الجدول رقم (8) يتضح أن بعض المعايير كانت متوفرة في ورقة الاختبار وبنسب تتراوح ما بين عالية ومتوسطة حيث كانت بين (96%) و(53%) وبشكل عام يمكن الحكم على أن المواصفات الفنية التي تم تحديدها في معيار وضع الاختبارات التحصيلية متوافرة بشكل جيد.

وفي هذه الدراسة حصل محور المعايير الفنية لإخراج الاختبار على نسبة استيفاء (82%) وهي نسبة عالية تدل على ممارسة عالية من قبل المعلمات وتتفق نتائج الدراسة مع دراسة (الكلثم، 2015)التي دلت على أن نسبة توافر مواصفات ورقة الاختبار الجيد تراوحت ما بين العالية والمتوسطة أما دراسات (العون، 2016) و(ساعد، 2013) دلت نتائجها على أن درجة إلمام المعلمين بمهارات إخراج الاختبار وكتابة التعليمات جاءت بدرجة متوسطة؛ وتخالف هذه الدراسة نتائج دراسة (المالكي، 2017) التي كان من أبرز نتائجها أنه هناك قصور في محور تعليمات الاختبار، ويعلل الباحثان ارتفاع هذا المحور لقيام وزارة التعليم بتوفير الطابعات الحديثة لضمان جودة الطباعة، كمان أنه يتم الزام المعلمات بكتابة التعليمات على ورقة الاختبار ويتم متابعة ذلك من قبل مديرة المدرسة والإشراف التربوي.

ب- ما مدى استيفاء اختبارات معلمات المرحلة المتوسطة لمعايير بناء فقرات الاختبار؟

نتائج تحليل الاختبار التحصيلي في ضوء معايير بناء الاختبارات لإخراج الاختبارات:

يوضح الجدول رقم (9) التالي مدي استيفاء الاختبارات المحللة لبنود المعيار في الجزء الخاص ببناء الفقرات (مواصفات نوع الأسئلة).

جدول رقم (9) مدى استيفاء الاختبارات المحللة للمعايير بناء الفقرات لإخراج الاختبار

م	الفقرات	مستوفي %	مستوفي إلى حد ما %	غير مستوفي%	الترتيب
	الأسئلة الموضوعية أراعي أن				
	أ_ اختيار من متعدد				
1	يطرح الجذر مشكلة واحدة.	%100.0	%0.0	%0.0	1
6	يقتصد في استخدام البديل (كل ما سبق) و (ليس واحد مما سبق).	%100.0	%0.0	%0.0	1
2	يصاغ الجذر على هيئة سؤال أو صيغة إكمال.	%93.3	%6.7	%0.0	2
4	تكون جميع البدائل متجانسة.	%93.3	%6.7	%0.0	2
8	أن توزع البدائل عشوائياً.	%90.0	%10.0	%0.0	3
7	لا يقل عدد البدائل عن أربع بدائل.	%83.3	0.0	%16.7	4
3	يظهر كل بديل في سطر مستقل إذا كان طويلاً.	%80.0	%0.0	%20.0	5
5	تكون البدائل قصيرة ما أمكن.	%50.0	%43.3	%6.7	6
	ب- أسئلة الصواب والخطأ				
2	تجنب كلمات التخصيص مثل فقط. في بعض الاحيان. لا يمكن	%94.4	%5.6	%0.0	1
5	احتواء العبارة على فكرة واحدة.	%94.4	%5.6	%0.0	1
1	تجنب التعميم مثل: كل. جميع. دائما. أحيانا. لا يمكن.	%83.3	%16.7	%0.0	2
4	يكون عدد الاجابات الصحيحة مساوي لعدد الاجابات الخاطئة ما أمكن.	%73.7	%15.8	%10.5	3
3	اتساق العبارات في الطول.	%16.7	%77.8	%5.6	4
	ج- أسئلة المزاوجة				
1	عدد بدائل القائمة الثانية أكثر من عدد العبارات في القائمة الأولى.	%84.6	%0.0	%15.4	1
2	تكون العبارات المستخدمة متجانسة.	%53.8	%7.7	%38.5	2
	د- فقرات الإكمال				
1	لا تحتوي عبارات الإكمال على أكثر من فراغين.	%68.8	%0.0	%31.3	1
2	تكون الفراغات عند نهاية العبارة أو قريب منها.	%68.8	%25.0	%6.3	1
	الأسئلة المقالية				
2	كل سؤال يقيس مهمة واحدة (لا يدمج من سؤالين او ثلاث في سؤال واحد).	%100.0	%0.0	%0.0	1
3	تحديد مسارات إجابة الطلبة من خلال تحديد نقاط أساسية في السؤال.	%100.0	%0.0	%0.0	1
1	تمثل الأسئلة جزء من الاختبار وليس كل الاختبار.	%92.3	%7.7	%0.0	2
	جميع معايير بناء الفقرات	%64.2	%31.0	%4.8	

في الجدول رقم (9) يتضح أن بعض المعايير كانت متوافرة في ورقة الاختبار وبنسب عالية إلى متوسطة تراوحت بين (100%) و(50%.) دلت نتائج الدراسة على أن نسبة ممارسة معلمات المرحلة المتوسطة لمعايير بناء الفقرات (64%) وهي تدل على ممارسه متوسطة.

ويرى الباحثان أن حصول هذا المحور على درجه متوسطة عائد إلى الاهتمام الذي توليه وزارة التعليم للاختبارات التحصيلية فبالإضافة إلى إقامة الدورات والورش التدريبية أيضاً يتم الاحتفاظ بنماذج للاختبارات في إدارات المدارس للسنوات السابقة ويسمح لمعلم المادة بالاطلاع عليها للاستفادة منها كما تتم مراقبة جودة الاختبارات

من قبل مكتب الاشراف التابع لوزارة التعليم. وبالرغم من ذلك فإن هذا المحور يعاني من بعض القصور في مجال بناء الأسئلة حيث يحتاج المعلمون إلى المزيد من الدورات والتدريب للوصول إلى درجه عالية من الممارسة.

ج- ما مدى التزام معلمات المرحلة المتوسطة بمعايير إخراج الاختبار؟

للإجابة على هذا السؤال تم استخدام النسب المئوية والنتائج مبينة في جدول رقم (10) والذي يبين نتائج تحليل عينة الدراسة في محاور الدراسة مرتبة من الأعلى استيفاء إلى الأقل حسب النسب المئوية كما يلي:

جدول (10) تحليل محاور الدراسة

	المحاور	نسبة الاستيفاء	مستوفي إلى حد ما	غير مستوفي	الترتيب
	محور المعايير الفنية	%82.2	%8.5	%9.3	1
	محور معايير بناء الفقرات	%64.2	%31.0	%4.8	3
	جميع محاور إخراج الاختبار	%69.8	%24.0	%6.2	2

- (المعايير الفنية) بنسبة استيفاء (82.2%) وقد احتل المرتبة الأولى.
- (معايير بناء الفقرات) بنسبة استيفاء (64.2%) وقد احتل المرتبة الثانية

وبصفة عامة يتبين أن النسبة المئوية لاستيفاء جميع المحاور تساوي (69.8%)؛ مما يدل على أن درجة ممارسة معلمات المرحلة المتوسطة لمعايير إخراج الاختبار في الاختبارات التحصيلية في منطقة تبوك متوسطة.

رابعاً: النتائج الخاصة بالسؤال الرابع: هل توجد فروق ذات دلالة إحصائية عند مستوى ($\alpha \leq 0.05$) في درجة التزام معلمات المرحلة المتوسطة بمعايير تخطيط وتصحيح الاختبار تُعزى إلى متغيري: الخبرة، والمؤهل العلمي والتفاعل بينهما؟

تم استخدام اختبار تحليل التباين الثنائي بين متوسطات إجابات المبحوثين في درجة ممارسة لمعرفة مدى التزام معلمات العلوم بشروط تخطيط وتصحيح الاختبار الجيد بتبوك تُعزى إلى متغير المؤهل العلمي والخبرة عند مستوى دلالة $\alpha \leq 0.05$ والنتائج مبينة في جدول رقم (11) و (12) كما يلي:

جدول (11) المتوسط الحسابي والانحراف المعياري لمهارتي التخطيط والتصحيح

المتغير	الفئة	مهارة التخطيط		مهارة التصحيح		المهارتان معاً	
		المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
المؤهل العلمي	بكالوريوس	4.32	0.43	4.29	0.43	4.30	0.394
	ماجستير	4.13	0.66	4.25	0.53	4.19	0.561
الخبرة	5 سنوات فأقل	4.24	0.48	4.23	0.50	4.23	0.458
	6- 10 سنوات	4.26	0.45	4.33	0.35	4.29	0.360
	أكثر من 10 سنوات	4.52	0.30	4.38	0.38	4.45	0.287

جدول (12) اختبار تحليل التباين الثنائي للكشف عن الفروق بين متوسطات درجة التزام معلمات المرحلة المتوسطة لمهارات التخطيط والتصحيح وفق متغيري المؤهل العلمي والخبرة.

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (F) المحسوبة	مستوى الدلالة sig
المؤهل	0.273	1	0.273	1.700	**0.194**
الخبرة	1.198	2	0.599	3.726	**0.026**
الخبرة *المؤهل	0.586	2	0.263	1.635	**0.198**
الخطأ	24.927	155	0.161		

يتضح من الجدول (12) ما يأتي:

1. لا توجد فروق بين متوسطات إجابات المبحوثين في درجة ممارسة معلمات المرحلة المتوسطة لمهارات التخطيط والتصحيح للاختبارات التحصيلية وفق معايير وشروط الاختبار الجيد تُعزى إلى متغير المؤهل العلمي. وقد يكون عدم وجود الفروق في متغير المؤهل عائد إلى جودة البرامج التربوية المقدمة لمرحلة البكالوريوس وغالباً ما تكون المعلمات من حملة البكالوريوس حاصلات على دبلوم في التربية إذ أنه من شروط تعيين خريجات الجامعات الاتي لم يدرسن مساقات تربوية في دراستهن الجامعية.

ونتيجة هذه الدراسة تتفق مع دراسة (الزهراني، 2011) ودراسة (الهاشمي، 2005) التي دلت على عدم وجود فروق تُعزى للمؤهل العلمي. وتختلف مع دراسة (العون، 2016) التي دلت على وجود فروق ذات دلالة إحصائية تُعزى إلى متغير المؤهل العلمي لصالح الدبلوم المتوسط ودراسة (الفرجات، 2015) التي أظهرت وجود فروق ذات دلالة إحصائية في كفايات المعلمين المتعلقة ببناء الاختبارات التحصيلية تبعا لمتغير المؤهل العلمي ولصالح المعلمين الحاصلين على المؤهل الأعلى، ودراسة (العمري، 2015) التي دلت على وجود فروق تُعزى لمتغير المؤهل في بناء الاختبارات التحصيلية.

2. توجد فروق بين متوسطات إجابات المبحوثين في درجة ممارسة معلمات المرحلة المتوسطة لمهارات التخطيط والتصحيح للاختبارات التحصيلية وفق معايير وشروط الاختبار الجيد بتبوك تُعزى إلى متغير الخبرة، وتتفق نتائج هذه الدراسة مع دراسة (المسعودين، 2017) ودراسة (الهاشمي، 2005) التي دلت على وجود فروق تُعزى لمتغير (الوظيفة، الجنس الخبرة) في مدى التزام المعلمين بمعايير الاختبارات التحصيلية؛ وتختلف نتائج هذه الدراسة مع دراسة (العون، 2016) ودراسة (الريامي، 2008)التي كان من نتائجها عدم وجود فروق ذات دلالة إحصائية تُعزى إلى متغير الخبرة مدى التزام المعلمين بمعايير الاختبارات التحصيلية.

3. لا توجد فروق بين متوسطات إجابات المبحوثين في درجة ممارسة معلمات المرحلة المتوسطة لمهارات التخطيط والتصحيح للاختبارات التحصيلية وفق معايير وشروط الاختبار الجيد بتبوك تُعزى إلى متغير المؤهل العلمي والخبرة معا (بمعنى أن تأثير المؤهل العلمي والخبرة معا لا يؤثران معنوياً على آراء أفراد العينة في درجة التزام معلمات المرحلة المتوسطة لمهارات التخطيط والتصحيح للاختبارات التحصيلية).

ولمعرفة مصدر دلالة الفروق؛ تم استخدام اختبار شفهي لتحديد الأوساط ذات الدلالة حسب متغير الخبرة.

جدول رقم (13) اختبار شفهي لتحديد الأوساط ذات الدلالة حسب متغير الخبرة

الفرق بين المتوسطات	5 سنوات فأقل	من 6- 10 سنوات	أكثر من 10 سنوات
5 سنوات فأقل		0.003	-0.396*
من 6- 10 سنوات	-0.003		-0.399*

الفرق بين المتوسطات	5 سنوات فأقل	من 6- 10 سنوات	أكثر من 10 سنوات
أكثر من 10 سنوات	0.396*	0.399*	

يتبين من اختبار شفيه جدول رقم (13) أن هناك فروق بين فئتي "أكثر من 10 سنوات"، "أقل من 5 سنوات" والفروق لصالح الفئة "أكثر من 10 سنوات"، كما توجد فروق بين فئتي "أكثر من 10 سنوات"، "5- 10 سنوات" والفروق لصالح الفئة "أكثر من 10 سنوات".

التوصيات والمقترحات

من خلال نتائج الدراسة الميداني يوصي الباحثان بما يلي:

1. عقد دورات تدريبية تستهدف تحسين كفايات المعلمات في مجال بناء الاختبارات التحصيلية الجيدة.
2. تشجيع المعلمات على تحكيم الاختبارات بعد كتابتها من قبل زميلاتهم المعلمات ذوات الخبرة للتأكد من ملاءمتها.
3. إنشاء قسم للقياس والتقويم بإدارة التعليم تكون من مهامه متابعة الاختبارات وإقامة الدورات ورفع مستوى المعرفة ببناء الاختبارات لدى المعلمات.
4. تكاتف جهود جميع الأطراف المعينة بالعملية التعليمية من أجل وضع قائمة معايير شاملة للاختبارات ككل، تكون بمثابة وثيقة يتم الرجوع إليها عند بناء الاختبارات.
5. ضرورة التكوين الجيد للمعلمات أثناء فترة تكوينهن، لاسيما في فترة دراستهن الجامعية، حول كيفية بناء الاختبارات التحصيلية الجيدة، ويتم التركيز على الجانب العملي التطبيقي أكثر من الجانب النظري.
6. وضع تصور مقترح لتطوير كفايات معلمات المرحلة المتوسطة في مجال إعداد الاختبارات التحصيلية الجيدة.
7. إجراء المزيد من الدراسات في مجال تقويم الاختبارات في المرحلة المتوسطة.

قائمة المراجع:

أولاً- المراجع بالعربية:

- ابن منظور، أبو الفضل جمال الدين محمد بن مكرم. (1993). لسان العرب. بيروت: دار الكتب.
- جابر، جابر عبد الحميد. (2002). اتجاهات وتجارب معاصرة في تقويم أداء التلميذ والمدرس. (ط1). القاهرة: دار الفكر العربي.
- الرويلي، مشعل مؤنس نخيلان. (2016). مهارات بناء الاختبارات التحصيلية لدى الطلاب المعلمين بالمملكة العربية السعودية: دراسة تحليلية تشخيصية من وجهة نظرهم والمختصين. مجلة التربية، جامعة الأزهر. مصر. ع170ج2: ص ص364-407.
- الريامي، حمد جمعة سليمان. (2008). درجة التزام معلمي الشهادة العامة بمعايير بناء الاختبارات التحصيلية المتبعة في سلطنة عمان. رسالة ماجستير، جامعة اليرموك، كلية التربية، الأردن.
- الزهراني، سعيد محمد سعيد. (2011). درجة امتلاك معلمي ومشرفي محافظة القريات لمبادئ بناء الاختبارات التحصيلية وممارستهم لها من وجهة نظرهم. رسالة ماجستير منشورة، جامعة اليرموك، كلية التربية، الأردن.
- ساعد، صباح. (2013). تقويم كفاية بناء الاختبارات التحصيلية لدى أساتذة التعليم الجامعي وفق معايير الاختبار الجيد. مجلة العلوم الإنسانية والاجتماعية. الجزائر. ع28: ص ص 81-90.

- سليمان، أمين علي محمد؛ وعلام، رجاء محمود. (2010). القياس والتقويم في العلوم الإنسانية أسسه وأدواته وتطبيقاته. (ط1). القاهرة: دار الكتاب الحديث.
- الشربيني، زكريا أحمد؛ وإبراهيم، رزق سند؛ ودمنهوري، رشاد صالح؛ ومطحنة، السيد خالد. (2010). القياس النفسي (ط1). الرياض: مكتبة الشقري.
- صنيدح، محمد فرحي. (2016). درجة تضمين كتب الاجتماعيات في المرحلة المتوسطة في دولة الكويت لمشروعات التعلم الخدمي. رسالة ماجستير غير منشورة، جامعة آل البيت. الأردن.
- الطراونة، عيسى عبد الوهاب سليم. (2012). تقويم الاختبارات التحصيلية بمدارس محافظة الكرك وفق معايير الاختبار الجيد، مجلة رابطة التربية الحديثة. مصر. مج 5، ع 15، ص ص: 183-228.
- عتوم، حمد صالح. (2014). تقويم الاختبارات التحصيلية من اعداد معلمي العلوم في محافظة جرش: دراسة تحليلية لنتائج الطلبة للفصل الدراسي الثاني للعام (2013/2012) في مبحث العلوم. مجلة جرش للبحوث والدراسات. الأردن. مج15، ع2: ص ص 361-375.
- عرفات، محمد جخراب. (2015). كفايات إعداد الاختبارات التحصيلية الجيدة لدى أساتذة التعليم الثانوي، مجلة العلوم الإنسانية، الجزائر. ع44: ص ص 515-540.
- العساف، صالح حمد. (2010). المدخل إلى البحث في العلوم السلوكية. (ط1). الرياض: دار الزهراء.
- علام، صلاح الدين. (2011). القياس والتقويم في العملية التدريسية. (ط5). القاهرة: دار الفكر العربي.
- عمارة، عبد الله سليم. (2008). المعايير التربوية والفنية لبناء أسئلة الاختبارات التحصيلية. مجلة التربية، قطر. س37، ع164: ص ص 150-171.
- العمري، ربيع محمود ضيف الله. (2015). مدى معرفة معلمي العلوم للمرحلة الأساسية في مديرية التربية والتعليم في لواء المزار الشمالي بمبادئ بناء الاختبارات التحصيلية. رسالة دكتوراه منشورة، جامعة أم درمان الإسلامية. معهد البحوث والدراسات الاستراتيجية. السودان.
- العنزي، يحيى محمد. (2017). أسباب انتشار العنف المدرسي لدى طلبة المرحلة المتوسطة من وجهة نظر المعلمين في مدارس محافظة القريات في المملكة العربية السعودية. رسالة ماجستير غير منشورة، جامعة اليرموك- إربد. الأردن.
- العون، إسماعيل سعود حنيان. (2016). مدى كفاءة مُعلمي ومعلمات التربية الرياضية في بناء الاختبارات التحصيلية في مدارس التربية والتعليم بالبادية الشمالية الشرقية الأردنية وفق معايير الاختبار الجيد. مجلة المنارة للبحوث والدراسات، الأردن. مج 17، ع2: ص ص 181-204.
- فتح الله. مندور عبد السلام. (2016). التقويم التربوي. (ط3). الرياض: دار النشر الدولي.
- الفرجات، هشام عقيلة علي. (2015). مدى كفاءة معلمي المدارس الحكومية الثانوية التابعة لمديرية التربية والتعليم بمحافظة العقبة في بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد. التربية، جامعة الأزهر. مصر. ع164: ص ص565-600.
- القاسم، محمد عبد الكريم محمود. (2016). تقويم الاختبارات الوزارية الموحدة في محافظات الضفة الغربية من وجهات نظر الإداريين والفنيين، رسالة ماجستير، كلية الدراسات العليا، جامعة النجاح الوطنية. نابلس.
- قدار، خالد محمد عباد. (2013). مدى توافر المعايير في الاختبارات التحصيلية الوزارية للمواد الاجتماعية للصف الثاني عشر الأدبي في الجمهورية اليمنية. التواصل. اليمن. ع31: ص ص 167-215.

- الكلثم، حمد مرضي إبراهيم. (2015). تقويم بعض اختبارات تخصص تربية إسلامية في جامعتي أم القرى والدمام في ضوء معايير جودة الاختبار. دراسات عربية في التربية وعلم النفس. السعودية. ع66: ص ص 211-236.
- المالكي، بدرية سعيد. (2017). دراسة تحليلية لمدى توافق الاختبارات الوطنية مع معايير الاختبار الجيد للمدارس المستقلة في المرحلة الابتدائية بدولة قطر. مجلة دراسات الخليج والجزيرة العربية. الكويت. س43، ع166: ص ص 21-61.
- مجمع اللغة العربية. (2010). المعجم الوسيط. القاهرة: دار الدعوة.
- مراد، صلاح أحمد؛ وسليمان، أمين علي. (2005). الاختبارات والمقاييس في العلوم النفسية والتربوية: خطوات إعدادها وخصائصها. (ط). القاهرة: دار الكتاب الحديث.
- المسعودين، أيمن عيد جراد. (2017). درجة التزام معلمي المرحلة الأساسية في الأردن بالمعايير الوطنية في بناء اختبارات التحصيلية من وجهة نظر المديرين والمشرفين التربويين. رسالة ماجستير منشورة. جامعة آل البيت. كلية العلوم التربوية. الأردن.
- مصبح، منور سالم إبراهيم. (2013). تقويم الاختبارات الموحدة لمبحث اللغة العربية في المرحلة الأساسية بمدارس وكالة الغوث بغزة وفق المعايير العالمية. رسالة ماجستير منشورة، الجامعة الإسلامية غزة، كلية التربية. فلسطين.
- النفيعي، عبد الرحمن عبد الله أحمد. (2016). دراسة تقويمية للتعرف على مدى إتقان أعضاء مشروع الاختبارات التحصيلية لمهارات بناء الاختبارات التحصيلية، العلوم التربوية مصر، مج 24، ع1: ص ص 215-253.
- الهاشمي، خليل ناصر سيف. (2005). مستوى معرفة المشرفين التربويين بسلطنة عمان بالمبادئ العامة للاختبارات التحصيلية. رسالة ماجستير غير منشورة، جامعة اليرموك. الأردن.
- يوسف، ماهر إسماعيل؛ والرفاعي، محب محمود. (2005). التقويم التربوي أسسه وإجراءاته. (ط4). الرياض: مكتبة الرشد.

ثانياً- المراجع الأجنبية:

- Bloom, S. (1976). Human Characteristics and School Learning. New York: Mc Graw- Hill.
- Gronlund, N. (1985). Measurement and evaluation in teaching (5th ed). London: Collier MacMillan Publishers.
- Martorella, p. (1991). Teaching social studies in middle and secondary school. Macmillan publishing company. New York.