

## تقييم جودة الاختبارات التحصيلية في المرحلة الثانوية في المملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين

Evaluation Of The Quality Of Achievement Tests In Secondary Education In  
The Kingdom In The Light Of Good Test Standards From The Point Of  
View Of Teachers

إعداد

**د. عيسى جود الله حميد الحربي**

أستاذ القياس والتقويم المشارك

الجامعة الإسلامية - المدينة المنورة - السعودية

**Dr. Essa J. H. Al-Harbi**

Associate Professor of Measurement and Evaluation

Islamic University of Madinah

Saudi Arabia

## تقييم جودة الاختبارات التحصيلية في المرحلة الثانوية في المملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين

إعداد

**د. عيسى جود الله حميد الحربي**

أستاذ القياس والتقويم المشارك

الجامعة الإسلامية - المدينة المنورة - السعودية

الملخص:

هدفت الدراسة الى الكشف عن مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة العربية السعودية في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين، واختلافها حسب النوع (معلم/معلمة)، والتخصص، وسنوات الخبرة، وتوصلت الدراسة إلى: الاختبارات التحصيلية في المرحلة الثانوية بشكل عام عالية الجودة. المجالات الأربعة (بنية الاختبار، تطبيق الاختبار، ورقة الاختبار، تصحيح الاختبار) عالية الجودة، بينما تم تقييم مجال (إحصائيات الاختبار) متوسط الجودة. تقييم المعلمات لجودة المجالات الثلاثة (بنية الاختبار، تطبيق الاختبار، ورقة الاختبار) والدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية أعلى بدلالة إحصائية من المعلمين. جودة (بنية الاختبار، إحصاءات الاختبار) من وجهة المعلمين ذوي التخصص (دراسات علمية) أعلى بدلالة إحصائية من ذوي التخصصات (دراسات أدبية، وحاسب وتقنية). في حين أن تقييم جودة (ورقة الاختبار) من وجهة المعلمين ذوي التخصص (دراسات أدبية) أعلى بدلالة إحصائية من تقييم التخصصات (دراسات علمية، وحاسب وتقنية). كلما ارتفع عدد سنوات الخبرة في التدريس ارتفع تقييم المعلمين لجودة الاختبارات التحصيلية. وقد أوصت الدراسة بالتالي:

حث المعلمين للالتزام بمعايير الاختبار الجيد عند إعداد الاختبارات التحصيلية.

عقد دورات للمعلمين لتدريبهم على بناء الاختبارات التحصيلية في ضوء معايير الاختبار الجيد.

عمل ندوات للمعلمين للاهتمام بتطوير الاختبارات التحصيلية ومطابقتها لمعايير الاختبار الجيد.

تطبيق الدراسة على عينات أخرى بمراحل التعليم المختلفة.

الكلمات المفتاحية : تقييم - جودة الاختبارات التحصيلية - معايير الاختبار الجيد - القياس والتقويم

## **Evaluation Of The Quality Of Achievement Tests In Secondary Education In The Kingdom In The Light Of Good Test Standards From The Point Of View Of Teachers**

**Dr. Essa J. H. Al-Harbi**

Associate Professor of Measurement and Evaluation

Islamic University of Madinah

Saudi Arabia

### **Abstract**

The study aimed to investigate the Saudi teachers' perceptions on the level of quality of achievement tests at the secondary schools stage of education in Saudi Arabia in the light of 'good test standards' (i.e., using the parameters of structure, application, test paper, correction and test statistics) from the point of view of teachers' perceptions. The research study also explored the difference of teachers' perceptions according to gender, area of specialization, and years of experience. The study found that the achievement tests at the Saudi secondary stage schools are generally of high quality in general. Results also indicated that female teachers' evaluation of the test quality is statistically significantly higher than that of the male teachers. Further, that teachers in the scientific area stream perceived the quality of (test structure, test statistics) as being statistically higher than those in the literary, computer and technical studies specializations. Moreover, teachers with specialization in literary studies had statistically higher perception of the test paper. Finally, the study showed that the higher the number of years of teaching experience, lead to the higher the teachers' evaluation to of the quality of achievement tests. The study recommended that teachers to adhere to good test standards when preparing achievement tests.

**Keywords:** Evaluation, tests quality, Achievement Tests, standards of good testing, measurement and evaluation

## مقدمة

تعمل الاختبارات التحصيلية على قياس وتقييم العملية التعليمية، والتي لا غنى عنها لمعرفة مقدار التقدم في تعلم وتحصيل الطلاب، ومدى تحقيقهم للأهداف التعليمية، مما أدى لبروز أهمية الاختبارات التحصيلية قديماً وحديثاً، وذلك كونها الأنسب في عملية قياس وتقييم المتعلمين من حيث تحديد درجة التغير لديهم في معارفهم ومهاراتهم وسلوكهم وقيمهم وذلك ضمن ضوابط واجراءات.

مما يؤكد على أن الاختبارات التحصيلية جزء لا يتجزأ من مكونات العملية التعليمية، بل إنها أداة مهمة للتعرف على مدى نجاح تنفيذ العملية التعليمية، وبالتالي العمل على تحسين وتطوير إجراءات آليات العملية التعليمية، كما أن الاختبارات تعد الإجراء الأكثر تداولاً لتقويم المتعلمين، والتعرف على مدى تقدمهم نحو الأهداف المنشودة (العابدي، 2006). وفي هذا الصدد أشار غرايسة وعواريب (2021) أن الاختبارات التحصيلية أحد الوسائل ذات الأهمية التي يتم الاعتماد عليها في عمليتي قياس وتقويم تقدم الطلاب في مراحلهم الدراسية المختلفة، كما يتم من خلالها الوقوف على نقاط القوة والضعف لدى الطلاب من أجل تطوير العملية التعليمية والارتقاء بها.

والاختبارات من أهم الأدوات التي تزود بالبيانات التي يعتمد عليها في اتخاذ القرارات المهمة كتصنيف الطلاب، وتحديد مساهمهم حسب قدراتهم ومهاراتهم، وتقييم تحصيلهم من خلال الدرجات التي يحصلون عليها في الاختبارات الفصلية والنهائية (عودة، 2000).

ويعد قياس مدى تحصيل الطلاب من المهمات التي تتميز بالصعوبة نظراً لكون التحصيل من السمات الكامنة التي يصعب قياسها بصورة مباشرة (أبو مديغم والعطاونة، 2021)، وبالتالي يرى العتوم (2014) أن الاختبارات التحصيلية كأداة للقياس تتطلب القدر الكافي من الاهتمام والتطوير كي تصبح عادلة وذات خصائص جيدة، وتتضح أهميتها كونها من المحكات التي يرجع إليها المعلمون في تقييم طلابهم، لذا يجب أن تبنى وتستخدم ضمن إجراءات تتصف بالضبط العلمي، وأن تراعي مواصفات ومعايير الاختبار الجيد من حيث الثبات والصدق والشمولية والوضوح، وان تراعي تصنيف بلوم بمجالاته الثلاثة المعرفي والوجداني والمهاري (الهويدي، 2004)؛ (نصار، 2009).

وتكتسب الاختبارات التحصيلية مزيداً من الأهمية لدورها في الارتقاء بالعملية التعليمية وتطويرها، لذا تم وضع مجموعة من المعايير لمواصفات الاختبار الجيد، وكشفت دراسة غرايسة وعواريب (2021) أن الاختبارات التحصيلية يجب أن تتوفر فيها معايير الجودة من حيث كتابة التعليمات، وجودة الطباعة، واكتمال فقرات الاختبار، في حين أكدت دراسة الغزال (2020) على ضرورة التزام الاختبارات التحصيلية بمعايير الجودة الشاملة، وأكدته أيضاً دراسة عثمان وعبدالله (2016) ودراسة أبو زيد والحسن (2017) أن الاختبارات التحصيلية المستخدمة يجب أن يُراعى في إعدادها مواصفات الاختبار الجيد، بينما كشفت دراسة بدر (2017) أن مستوى جودة الاختبارات التحصيلية كان منخفض من حيث بناء الاختبار وشموليته، وخلوه من التعليمات، وإخراج الاختبار، وطباعته.

وعن أهم صفات تصميم الاختبار التحصيلي الجيد، أوضحت دراسة محمد العبيدي (2011) أنها تتمثل في الصدق، والثبات، والموضوعية، والشمولية، والتميز.

أما عن فاعلية الاختبارات التحصيلية والتي تعزى لتغير المؤهل العلمي أو الخبرة أو النوع الاجتماعي ووجهات نظر المعلمين، فقد توصلت دراسة أبو مديغم والعطاونة (2021) إلى عدم وجود فروق من وجهة نظر المعلمين حول فاعلية الاختبارات التحصيلية في ضوء معايير بناء الاختبار الجيد تعزى إلى هذه المتغيرات، وأكدت أنه لا يوجد فروق في كفايات بناء الاختبارات التحصيلية لدى المعلمين ترجع إلى الجنس أو مستوى الخبرة، وفي دراسة الجلندانية وآخرون (2022) تبين أن مستوى الرضا لدى أعضاء هيئة التدريس عن الاختبارات منخفض، أما الطلاب فكان متوسط، وأنه لا يوجد فروق في مستوى الرضا تعزى إلى التخصص، وكشفت دراسة فرطاس (2021) أن أعضاء هيئة التدريس يلتزمون بدرجة كبيرة بتطبيق معايير جودة الاختبارات بدرجة كبيرة، ولا توجد فروق في تطبيق معايير جودة الاختبارات تعزى إلى اختلاف التخصص، في حين أشارت دراسة النوافلة وآخرون (2021) إلى أن اتجاهات المعلمين نحو الاختبارات التحصيلية متوسطة، ولا توجد فروق في اتجاهاتهم تعزى لاختلاف الجنس أو الخبرة، وفي دراسة الزبون (2017) اتضح أن درجة معرفة المعلمين في مجال الاختبارات التحصيلية كانت متدنية، ودرجة ممارستهم كانت متوسطة، ومعوقات بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد كانت متوسطة. ويتفق ذلك مع ما أكدته دراسة إبراهيم (2016) من حيث وجود ضعف كبير لدى المعلمين في بناء الاختبارات التحصيلية وإعدادها وفق معايير الاختبار الجيد.

وعلى مستوى الطلاب أكد الرويلي (2016) أن معرفة الطلاب بمهارات القياس والتقويم الأساسية لبناء الاختبارات التحصيلية متدنية، ودرجة إتقانهم وتركيزهم على هذه المهارات متوسطة، وفي إطار كفايات بناء الاختبارات التحصيلية الجيدة فقد أكدت دراسة عدائكة وآخرون (2018) ودراسة ضيف والأسود (2021) أن مستوى كفايات إعداد الاختبارات التحصيلية لدى المعلمين منخفضة، وتوصلت دراسة عزي وآخرون (2018) إلى أن المعلمين لديهم احتياجات تدريبية في كيفية بناء الاختبارات التحصيلية وكذلك تحليلها وتفسير نتائج التلاميذ في ضوءها.

### أهداف الاختبارات التحصيلية:

يُعد الهدف الرئيسي من تصميم الاختبارات التحصيلية هو قياس مدى تعلم الطلاب للمفاهيم والمعارف والمهارات المتعلقة بالمادة التعليمية خلال مدة، أو نهاية فترة دراسية، أو وفقاً لمعيار أو شرط معين، وتتمثل أهداف الاختبارات التحصيلية في التوصل إلى:

- قدرات الطلاب ومدى كفاءتهم.
- درجة التقدم في تعلم الطالب.
- احتياجات الطلاب التعليمية.

- جوانب القوة والتميز والضعف.
- القرارات التعليمية والتحصيلية.
- تقييم المكونات التعليمية.

#### مواصفات الاختبار الجيد:

- ينبغي أن يتصف الاختبار بدرجة عالية من:
- الصدق: أن يقيس الاختبار ما وضع لأجله.
- الثبات: أن يعطي الاختبار نتائج متقاربة عند إعادة تطبيقه على نفس المختبرين.
- الموضوعية: أن يتم الالتزام بالموضوعية في بناء الاختبار والتعامل معه وتصحيحه أو تفسير نتائجه.
- التمييز: قدرة الاختبار على كشف التفاوت والتباين في موضوع الاختبار بين المختبرين.
- الشمولية: أن يغطي الاختبار بشكل يعكس حجم أهداف ومكونات موضوع الاختبار.

#### خطوات بناء الاختبارات التحصيلية:

يخضع بناء الاختبارات التحصيلية لعدد من الخطوات التي تضمن تحقق مواصفات الاختبار الجيد، والتي تتمثل

في:

- تحديد الهدف والغاية من الاختبار.
- تحديد محتوى المقرر الدراسي المقاس.
- تحديد مكونات المحتوى الدراسي وإعطاء كل مكون وزن نسبي.
- تحديد الأهداف التعليمية المقاسة وإعطاء كل هدف وزن نسبي.
- تحديد عدد الأسئلة وتوزيعها وفقاً للنسب مع ضرورة تقدير الوقت المتاح للاختبار.
- إعداد وتصميم جدول المواصفات.
- اختيار نوع أسئلة الاختبار وصياغتها.
- تحكيم الأسئلة من قبل الزملاء والخبراء مع مراجعة الأوزان.
- إعداد نموذج إجابة ووصف آلية توزيع الدرجات وتفسيرها.
- صياغة تعليمات وضوابط تنفيذ الاختبار.
- إخراج الاختبار واعتماده.

## أنواع الاختبارات التحصيلية:

تتعدد معايير تصنيف الاختبارات التحصيلية، ويمكن حصرها في ثلاثة مجالات، وذلك من حيث:

### 1- الهدف: وفيه يتم تصنيف الاختبارات التحصيلية إلى:

❖ اختبارات مسحية.

❖ اختبارات تحديد مستوى.

❖ اختبارات تشخيصية.

### 2- التقنين: وفيه يتم تصنيف الاختبارات التحصيلية إلى:

❖ اختبارات مقننة.

❖ اختبارات غير مقننة.

### 3- الاستجابة: وفيه يتم تصنيف الاختبارات التحصيلية إلى:

❖ اختبارات شفوية.

❖ اختبارات تحريرية.

❖ اختبارات أداء.

## مشكلة الدراسة:

في ضوء أهمية الاختبارات التحصيلية، وبناءً على ما تم ملاحظته من تفاوت في قدرات المعلمين في إعداد الاختبارات التحصيلية في ضوء معايير الاختبار الجيد، تتمثل مشكلة الدراسة في ضرورة تقييم جودة الاختبارات التحصيلية في ضوء معايير الاختبارات التحصيلية، حيث تتحدد مشكلة الدراسة الحالية في الإجابة على الأسئلة الآتية:

ما مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين؟

هل يختلف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف النوع (معلم/معلمة) من وجهة نظر المعلمين؟

هل يختلف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف التخصص من وجهة نظر المعلمين؟

هل يختلف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف سنوات الخبرة في التدريس من وجهة نظر المعلمين؟

## أهداف الدراسة:

1. تهدف الدراسة إلى تقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين.
2. دراسة اختلاف جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف كل من: النوع (معلم/معلمة)، والتخصص، وسنوات الخبرة في التدريس من وجهة نظر المعلمين.

## أهمية الدراسة:

1. تنبع أهمية الدراسة من أهمية الاختبارات التحصيلية في عمليتي قياس وتقويم مستويات وقدرات الطلاب.
2. تعد هذه الدراسة من الدراسات التي تناولت جودة الاختبارات التحصيلية.
3. تساعد الدراسة القيادات التربوية في الميدان التعليمي على الاهتمام بالتنمية المهنية للمعلمين في المرحلة الثانوية وسائر المراحل التعليمية في تصميم وإعداد الاختبارات التحصيلية بما يتوافق مع معايير الاختبار الجيد.

## مصطلحات الدراسة:

**الاختبارات التحصيلية:** أداة قياس تعد بشكل يتصف بالضبط والتنظيم وإجراءات مقيدة بشروط وقواعد تهدف لتحديد درجة تواجد خاصية أو سمة أو قدرة ما لدى المفحوص من خلال استجابته لعينة من المثيرات التي تمثل الخاصية المقاسة أو السمة أو القدرة.

فيما عرفها (اليقوي، 2013) بأنها الأداة التي تستخدم في قياس معرفة وفهم ومهارة الطالب في مادة دراسية أو تدريبية معينة، و(الزهراني، 2009) بأنها أداة قياس يتم إعدادها وفقاً لطريقة منظمة من عدة خطوات تتضمن مجموعة من الإجراءات التي تخضع لشروط وقواعد محددة بغرض تحديد درجة امتلاك سمة أو قدرة معينة لدى الفرد من خلال الاستجابة لعينة من المحفزات التي تمثل السمة أو القدرة المرغوبين قياسهم، وعرفها جاسم (2011) بأنها وسيلة تستخدم لقياس مستوى وخبرات الطلاب في مقرراتهم الدراسية، والتي تفيد في تحديد ترتيب الطلاب ومراكزهم في خبرة معينة مقارنة بالمجموعة التي ينتمون إليها، وعرفها الغزال (2020) بأنها مجموعة من الأسئلة والمواقف يطلب من الطالب الإجابة عليها، على أن يراعى في تصميمها الأسس والمعايير التربوية الخاصة بإعدادها وتصميمها وإخراجها بما يتوافق مع معايير الجودة الشاملة، أما محمد (2021) فقد عرفها بأنها وسيلة يستخدمها المعلمون والتي تتكون من مجموعة من المثيرات التي تستدعي مجموعة من استجابات الطلاب انطلاقاً مما اكتسبوه من معلومات ومعارف ومهارات سواء أكانت هذه المهارات أكاديمية أو وجدانية أو بدنية.



**معايير الاختبار الجيد:** وهي متطلبات أساسية للاختبارات يتم الحكم على جودة الاختبار تبعاً لدرجة تواجد كل متطلب.

وقد عرفها (عبد الواحد، 2015) بأنها مجموعة من المحددات الأساسية ذات فاعلية تستخدم في الحكم على جودة الاختبار التحصيلي من حيث محتواه وصياغته، كما عرفت في الحري (2007) والدوسري (2004) بأنها عدد من الخصائص السيكومترية، والاشتراطات، والمواصفات، وتحليل المحتوى، واشتقاق الأهداف، وصياغة عبارات الاختبار، وتحليله، وتشخيصه، واتخاذ القرارات التربوية في ضوءه.

#### حدود الدراسة:

**الحدود الموضوعية:** تناولت الدراسة تقييم جودة الاختبارات التحصيلية في المرحلة الثانوية في المملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين.  
**الحدود الزمانية:** العام الجامعي 1442هـ.  
**الحدود المكانية:** مدارس المرحلة الثانوية بالمملكة العربية السعودية.

#### منهج وإجراءات الدراسة

##### منهج الدراسة

في ضوء مشكلة وأهداف الدراسة، تم اللجوء إلى المنهج الوصفي، لوصف تقييم جودة الاختبارات التحصيلية في المرحلة الثانوية في المملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين.

#### عينة الدراسة:

تكونت عينة الدراسة من (607) من معلمي ومعلمات المرحلة الثانوية، ويوضح الجدول التالي توزيع عينة الدراسة حسب المتغيرات الثلاثة: (النوع، الخبرة، والتخصص):

جدول (1) توزيع عينة الدراسة حسب المتغيرات الثلاثة: (النوع، الخبرة، والتخصص)

		معلمون			معلمات				
الخبرة	5	10-6	11-	أكثر	5	6-10	11-	أكثر	
	سنوات	سنوات	سنة	مجموع	سنوات	سنوات	سنة	مجموع	
التخصص	أقل	سنوات	سنة	15	أقل	سنوات	سنة	15	
	دراسات	أدبية	16	54	28	19	117	23	
		21	19	37	21	19	37	23	
		100	21	19	100	21	19	217	

دراسات علمية	26	66	26	23	141	31	44	22	18	115	256
حاسب وتقنية	19	27	15	11	72	10	18	25	9	62	134
المجموع	61	147	69	53	330	64	99	66	48	277	607

## أداة الدراسة:

### مقياس جودة الاختبارات التحصيلية:

أعدّه الباحث مقياس جودة الاختبارات التحصيلية، حيث يتألف من 22 عبارة أو بند، يتم الإجابة عليها باختيار (تدرّج نسبي) موزع على خمسة بدائل تدل على درجة تحقق العبارة: 0%، 25%، 50%، 75%، 100%، وهذه العبارات تقيس خمسة مجالات وهي:

1. بنية الاختبار: ويتكون من 4 عبارات.
2. إحصاءات الاختبار: ويتكون من 5 عبارات.
3. تطبيق الاختبار: ويتكون من 4 عبارات.
4. ورقة الاختبار: ويتكون من 5 عبارات.
5. تصحيح الاختبار: ويتكون من 4 عبارات.

وقد تم حساب معاملات ثبات وصدق مقياس جودة الاختبارات التحصيلية، وذلك بعد تطبيقه على عينة استطلاعية تكونت من (120) معلّمًا ومعلّمة تم اختيارهم بشكل عشوائي من العينة الأساسية للدراسة، حيث تم التوصل للثبات والصدق كما يلي:

### (1) ثبات المقياس:

تم حساب ثبات عبارات مقياس جودة الاختبارات التحصيلية بطريقتين هما:

- معامل ألفا ل كرونباخ Alpha-Cronbach (بعدد عبارات كل مجال)، وفي كل مرة يتم حذف درجات إحدى العبارات من الدرجة الكلية للمجال الذي تقيسه العبارة.
- حساب معاملات الارتباط بين درجات العبارة والدرجات الكلية للمجال الذي يقيسه.
- وتم حساب الثبات الكلي للمجالات وللمقياس ككل بطريقتي ألفا ل كرونباخ والتجزئة النصفية لـ Spearman-Brown

### (2) صدق المقياس:

- تم حساب صدق عبارات المقياس عن طريق حساب معامل الارتباط بين درجة العبارة والدرجة الكلية للمجال الذي يقيسه عند حذف درجة العبارة من المجال باعتبار أن بقية عبارات المجال محكًا لصدق العبارة.

- تم حساب صدق مجالات المقياس عن طريق حساب معامل الارتباط بين درجة المجال والدرجة الكلية للمقياس، فكانت النتائج كما بالجدول الآتي:

جدول (2) معاملات ثبات وصدق عبارات مقياس جودة الاختبارات التحصيلية (ن=120)

المجال	العبرة	معامل ألفا ل كرونباخ Alpha- Cronbach	معامل ارتباط درجة العبرة بالدرجة الكلية للمجال (ثبات)	معامل ارتباط درجة العبرة بالدرجة الكلية للمجال عند حذف درجة العبرة من المجال (صدق)
بنية الاختبار	1	0.646	**0.72	**0.48
معامل ألفا الكلي للمجال = 0.705	2	0.652	**0.72	**0.47
3	0.611	**0.76	**0.54	
معامل الثبات بطريقة التجزئة النصفية = 0.737	4	0.658	**0.71	**0.46
إحصاءات الاختبار	5	0.653	**0.76	**0.55
معامل ألفا الكلي للمجال = 0.724	6	0.670	**0.74	**0.50
7	0.657	**0.76	**0.54	
معامل الثبات بطريقة التجزئة النصفية = 0.771	8	0.700	**0.65	**0.42
9	0.705	**0.41	**0.41	
تطبيق الاختبار	10	0.689	**0.69	**0.44
معامل ألفا الكلي للمجال = 0.716	11	0.610	**0.77	**0.58
12	0.583	**0.80	**0.61	
معامل الثبات بطريقة التجزئة النصفية = 0.761	13	0.720	**0.68	**0.40
ورقة الاختبار	14	0.721	**0.73	**0.53
معامل ألفا الكلي للمجال = 0.763	15	0.735	**0.71	**0.50
16	0.697	**0.76	**0.60	
معامل الثبات بطريقة التجزئة النصفية = 0.769	17	0.725	**0.70	**0.52
18	0.726	**0.70	**0.52	
ورقة الاختبار	19	0.701	**0.69	**0.43
معامل ألفا الكلي للمجال = 0.717	20	0.601	**0.79	**0.59
21	0.629	**0.77	**0.55	
معامل الثبات بطريقة التجزئة النصفية = 0.750	22	0.684	**0.70	**0.46

معامل ألفا ل	معامل ارتباط درجة	معامل ارتباط درجة العبارة	المجال	العبارة
كرونباخ	العبارة بالدرجة الكلية للمجال	بالدرجة الكلية للمجال عند حذف درجة العبارة من المجال (صدق)		
Alpha-Cronbach	الكلية للمجال (ثبات)			
معامل Alpha-Cronbach العام الكلي للمقياس = 0.904				
معامل الثبات الكلي للمقياس بطريقة التجزئة النصفية ل Spearman-Brown = 0.912				

\*\* دال إحصائيًا عند مستوى (0.01)

يتضح من الجدول السابق ما يأتي:

- أن جميع معاملات ألفا لـ كرونباخ للمجال في حالة غياب أي عبارة من عباراته أقل من أو يساوي معامل ألفا للمجال في حالة وجود جميع عباراته، أي أن وجود عبارات أي مجال يساعد في ارتفاع معامل الثبات الكلي للمجال، وهذا يشير إلى أن كل عبارة تسهم بدرجة معقولة في الثبات الكلي للمجال الذي يقيسه.
- أن جميع معاملات ارتباط درجة العبارة بالدرجة الكلية للمجال الذي يقيسه دالة إحصائيًا عند مستوى (0.01)، مما يدل على الاتساق الداخلي وثبات جميع عبارات مقياس جودة الاختبارات التحصيلية.
- أن معاملات الثبات الكلي سواء للمجالات الفرعية أو للمقياس ككل بطريقتي Alpha-Cronbach، التجزئة النصفية لـ Spearman-Brown معاملات مرتفعة مما يدل على الثبات الكلي للمجالات الفرعية وللمقياس ككل.
- أن جميع معاملات ارتباط درجة العبارة بالدرجة الكلية للمجال الذي يقيسه (في حالة حذف درجة العبارة من الدرجة الكلية للمجال) دالة إحصائيًا عند مستوى (0.01)، مما يدل على صدق جميع عبارات مقياس جودة الاختبارات التحصيلية.
- وقد بلغت معاملات الارتباط بين درجة المجالات الخمسة والدرجة الكلية للمقياس على حسب ترتيب المجالات في الجدول السابق: (0.76)، (0.80)، (0.79)، (0.81)، (0.80) وهي معاملات مرتفعة ودالة إحصائيًا عند مستوى (0.01) مما يدل على صدق جميع مجالات مقياس جودة الاختبارات التحصيلية.
- من الإجراءات السابقة تأكد للباحث ثبات وصدق مقياس جودة الاختبارات التحصيلية، ومن ثم صلاحيته لقياس جودة الاختبارات التحصيلية في المرحلة الثانوية، حيث تشير الدرجة العالية على المقياس إلى ارتفاع جودة الاختبارات التحصيلية، بينما تشير الدرجة المنخفضة إلى انخفاض جودة الاختبارات التحصيلية.

#### الأساليب الإحصائية:

تم استخدام مجموعة من الأساليب الإحصائية للإجابة على أسئلة الدراسة وهي:

- المتوسط الحسابي والتكرارات والنسبة المئوية.
- اختبار (ت) T-test للعينتين المستقلتين.

- تحليل التباين متعدد المتغيرات التابعة (MANOVA) تحليل التباين متعدد المتغيرات التابعة (MANOVA) متبوعاً باختبار أقل فرق دال (Least significant difference) LSD للمقارنات المتعددة.

### عرض ومناقشة النتائج

نظراً لأن مدى الدرجات على عبارات المقياس يمتد من (1) إلى (5)، فقد تم استخدام الأسلوب التالي لتحديد مستوى الاستجابة على مقياس جودة الاختبارات التحصيلية، حيث تم تحويل درجة كل معلم أو معلمة على المقياس إلى 5 درجات وذلك عن طريق القسمة على عدد عبارات المقياس وهو 22، وبعد ذلك تم تصنيف تلك الإجابات إلى (5) مستويات أو فئات متساوية في المدى من خلال المعادلة التالية:

$$\text{طول الفئة} = (\text{أكبر وزن} - \text{أقل وزن}) \div \text{عدد بدائل المقياس} = (5-1) \div 5 = 0.80$$

بحيث يكون التصنيف الموضح بالجدول التالي:

جدول (3) توزيع مدى الاستجابة وفق التدرج المستخدم في المقياس

م	الاستجابة	مدى الاستجابة	درجة الجودة
1	100%	من 4.20 إلى 5	عالية جداً
2	75%	من 3.40 لأقل من 4.20	عالية
3	50%	من 2.60 لأقل من 3.40	متوسطة
4	25%	من 1.80 لأقل من 2.60	منخفضة
5	0%	من 1 لأقل من 1.80	منخفضة جداً

وفيما يلي نتائج الإجابة على أسئلة الدراسة:

### السؤال الأول:

للإجابة عن السؤال الأول ما مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين؟ تم استخدام:

- المتوسط الحسابي والنسبة المئوية للمتوسط بالنسبة للدرجة الكلية.
- التكرارات والنسبة المئوية لعدد المعلمين ذوي مستويات التقييم المختلفة.
- المتوسط الحسابي لكل عبارة من عبارات المقياس.

فكانت النتائج كما بالجدولين التاليين:

جدول (4) متوسط الدرجات والتكرارات والنسبة المئوية لعدد المعلمين ذوي مستويات التقييم المختلفة على (مقياس جودة الاختبارات التحصيلية) والنسبة المئوية للمتوسط (ن = 607)

مستويات جودة الاختبارات التحصيلية				
100%	75%	50%	25%	0%
عالية جداً	عالية	متوسطة	منخفضة	منخفضة جداً

من 1 لأقل من 1.80		من 1.80 لأقل من 2.60		من 2.60 لأقل من 3.40		من 3.40 لأقل من 4.20		من 4.20 إلى 5	
النسبة	العدد	النسبة	العدد	النسبة	العدد	النسبة	العدد	النسبة	العدد
-	-	-	-	8.7%	53	90.1%	547	1.2%	7
<b>المتوسط الكلي للمقياس * = 3.68</b>									
<b>النسبة المئوية للمتوسط الكلي للمقياس = 73.6%</b>									

\* المتوسط الكلي للمقياس من 5 درجات.

يتضح من الجدول السابق ما يلي:

- أن متوسط تقييم (جودة الاختبارات التحصيلية) في المرحلة الثانوية بالمملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين قد بلغ (3.68 من 5) ونسبة هذا المتوسط 73.6% من الدرجة العظمى للمقياس، كما أن هذا المتوسط يقع في مدى الاستجابة عالية (الذي يمتد من 3.40 لأقل من 4.20). مما يشير إلى أن معلمي ومعلمات المرحلة الثانوية بالمملكة بعينة الدراسة يُقِيمون بوجه عام الاختبارات التحصيلية في المرحلة الثانوية على أنها عالية الجودة.
- أن معظم معلمي ومعلمات المرحلة الثانوية بالمملكة ب عينة الدراسة يُقِيمون الاختبارات التحصيلية في المرحلة الثانوية على أنها عالية الجودة حيث بلغت نسبتهم (90.1%) من إجمالي عينة الدراسة. أما نسبة المعلمين الذين قِيموا الاختبارات التحصيلية في المرحلة الثانوية على أنها ذات جودة عالية جداً أو ذات جودة متوسطة نسب ضئيلة، حيث بلغت نسبتهم (1.2%)، (8.7%) على الترتيب من إجمالي عينة الدراسة، وأنه لم يُقِيم أي معلم أو معلمة الاختبارات التحصيلية في المرحلة الثانوية على أنها ذات جودة منخفضة أو ذات جودة منخفضة جداً.

جدول (5) المتوسط الحسابي والانحراف المعياري وترتيب كل عبارة بالمقياس (ن=607)

م	المجال	العبارة	المتوسط من (5)	الترتيب	درجة الجودة
1	بنية الاختبار	يغطي الاختبار جميع عناصر المحتوى وجميع مستويات الأهداف	3.77	7	عالية
2		يتقيد بالموضوعات الأساسية للمحتوى الدراسي	3.71	10	عالية
3	(3.70)	يراعى الاختبار جدول المواصفات والوزن النسبي للوحدات الدراسية	3.72	9	عالية
4		الاختبار قابل للتطبيق وفقاً للزمن المخصص للاختبار	3.61	14	عالية
5	إحصاءات الاختبار	الاختبار قادر على إبراز الفروق الفردية بين الطلاب	3.34	17	متوسطة
6		الاختبار يؤدي إلى توزيع طبيعي لدرجات المختبرين	3.35	16	متوسطة
7	(3.33)	يتصف الاختبار بالصدق أي أنه يقيس ما وضع لقياسه	3.32	18	متوسطة
8		الاختبار يتصف بالثبات عند إعادة الاختبار بنفس الظروف	3.39	15	متوسطة

متوسطة	19	3.27	فقرات الاختبار تتوزع بين السهولة والصعوبة	9
عالية	12	3.67	بيئة الاختبار ومكانه مجهزة ومهيأة	10
عالية	7	3.77	تعليمات الجلوس والتحرك ومواعيدها واضحة للمختبر	11
عالية	6	3.80	توزيع الأوراق يتم بشكل سلس ومنظم	12
عالية	8	3.74	الدعم المقدم للمختبرين إرشادي ولا يؤثر على أدائهم	13 (3.75)
عالية	10	3.71	تحتوي ورقة الاسئلة على المعلومات الشخصية للمختبر ومعلومات الاختبار	14
عالية	13	3.65	طريقة توزيع وإعطاء الدرجات واضحة للمختبر	15 ورقة
عالية	11	3.69	ورقة الاختبار تتصف بوضوح الطباعة وجمال التنسيق	16 الاختبار
عالية	5	3.82	صياغة الأسئلة سليمة نحويًا وخالية من الأخطاء الإملائية	17 (3.74)
عالية	4	3.83	ورقة الأسئلة منظمة ومنسقة الإخراج	18
عالية	3	3.92	طريقة التصحيح موحدة وواضحة	19
عالية	2	3.96	يتم التصحيح وفقا لمفتاح ومعايير محددة	20 تصحيح
عالية	2	3.96	المختبرين يدركون آلية توزيع وإعطاء الدرجات	21 الاختبار
عالية	1	4.02	يتم تدقيق ومراجعة التصحيح	22 (3.96)

يتضح من الجدول السابق ما يلي:

- تكوّن مقياس (جودة الاختبارات التحصيلية) من 22 عبارة، جاء منهم 17 عبارةً بمستوى جودة عالية، 5 عبارات بمستوى جودة متوسطة، ولا يوجد أي عبارة بمستوى عالية جداً أو ذات جودة منخفضة أو منخفضة جداً.
- أن أعلى عبارات مقياس (جودة الاختبارات التحصيلية) من حيث درجة الجودة هو العبارة (يتم تدقيق ومراجعة التصحيح)؛ يليه العبارتان (يتم التصحيح وفقاً لمفتاح ومعايير محددة)، (المختبرين يدركون آلية توزيع وإعطاء الدرجات) وفي المرتبة الثالثة العبارة (طريقة التصحيح موحدة وواضحة) حيث بلغ المتوسط الحسابي لهذه العبارات (4.02)، (3.96)، (3.96)، (3.92) على الترتيب وجميع هذه المتوسطات تشير إلى درجة جودة عالية. وأن أدناها هو العبارة: (فقرات الاختبار تتوزع بين السهولة والصعوبة)؛ حيث بلغ المتوسط الحسابي (3.27) وهو متوسط يشير إلى درجة جودة متوسطة. وأن الفرق بين أعلى متوسط وأقل متوسط هو  $(4.02 - 3.27) = 0.75$ ، وهو فرق ليس كبيراً، مما يدل على وجود تقارب في درجة جودة جميع عبارات مقياس (جودة الاختبارات التحصيلية) في المرحلة الثانوية من وجهة نظر معلمي ومعلمات المرحلة الثانوية.
- أما على مستوى المجالات، يلاحظ أن مجال (تصحيح الاختبار) احتل المرتبة الأولى بمتوسط تقييم (3.96) من (5) يليه مجال (تطبيق الاختبار) بمتوسط (3.75) وفي المرتبة الثالثة مجال (ورقة الاختبار) بمتوسط (3.74)، ثم مجال (بنية الاختبار) بمتوسط (3.70) وجميع هذه المتوسطات تشير إلى درجة جودة عالية لهذه المجالات الأربعة،

واحتل المرتبة الخامسة والأخيرة تقييم مجال (إحصاءات الاختبار) بمتوسط (3.33) وهو متوسط تقييم يشير إلى درجة جودة متوسطة.

### السؤال الثاني:

للإجابة عن السؤال الثاني هل يختلف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف النوع (معلم/معلمة)؟ تم استخدام:

- اختبار (ت) T-test للعينتين المستقلتين، والجدول الآتي يوضح نتائج هذا السؤال:

جدول (6) نتائج اختبار (ت) لدلالة الفرق بين متوسطي تقييم المعلمين والمعلمات لجودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة

م	المجال	المعلمون (ن=330)		المعلمات (ن=277)		قيمة (ت) ودالاتها
		المتوسط	الانحراف المعياري	المتوسط	الانحراف المعياري	
1	بنية الاختبار	3.62	0.47	3.81	0.46	**5.05
2	إحصاءات الاختبار	3.34	0.41	3.33	0.43	0.19 غير دالة
3	تطبيق الاختبار	3.63	0.49	3.88	0.43	**6.47
4	ورقة الاختبار	3.69	0.50	3.80	0.55	**2.77
5	تصحيح الاختبار	3.96	0.46	3.97	0.42	0.24 غير دالة
	الدرجة الكلية لجودة الاختبارات	3.63	0.23	3.74	0.22	**5.76

\*\* دال إحصائياً عند مستوى (0.01)

يتضح من الجدول السابق ما يلي:

- وجود فرق دال إحصائياً عند مستوي (0.01) بين متوسطي تقييم المعلمين والمعلمات لجودة المجالات الثلاثة (بنية الاختبار، تطبيق الاختبار، ورقة الاختبار) والدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة لصالح متوسط تقييم المعلمات في جميع الحالات. أي أن متوسط تقييم المعلمات لجودة المجالات الثلاثة (بنية الاختبار، تطبيق الاختبار، ورقة الاختبار) والدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة أعلى بدلالة إحصائية من نظائرها لدى المعلمين.
- عدم وجود فرق دال إحصائياً بين متوسطي تقييم المعلمين والمعلمات لجودة المجالين (إحصاءات الاختبار، تصحيح الاختبار). أي أنه يوجد تقارب بين متوسطي تقييم المعلمين والمعلمات لجودة مجالي الاختبارات التحصيلية (إحصاءات الاختبار، تصحيح الاختبار) في المرحلة الثانوية بالمملكة.

### السؤال الثالث:



للإجابة عن السؤال الثالث هل يختلف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف التخصص من وجهة نظر المعلمين؟ تم استخدام:

• تحليل التباين متعدد المتغيرات التابعة Multivariate Analysis of Variance MANOVA متبوعاً باختبار أقل فرق دال (Least significant difference) LSD للمقارنات المتعددة، ونتائج هذا السؤال بالجدولين التاليين:

جدول (7) نتائج تحليل التباين متعدد المتغيرات التابعة عند دراسة اختلاف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف التخصص (ن=607)

مصدر التباين	المجال	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	مستوى الدلالة
التخصص	بنية الاختبار	1.93	2	0.96	4.36	0.01
	إحصاءات الاختبار	1.87	2	0.94	5.44	0.01
	تطبيق الاختبار	1.67	2	0.83	3.60	0.05
	ورقة الاختبار	2.47	2	1.24	4.55	0.01
	تصحيح الاختبار	0.26	2	0.13	0.67	0.51
	الدرجة الكلية للتقييم	0.20	2	0.10	1.88	0.15
الخطأ	بنية الاختبار	133.61	604	0.22		
	إحصاءات الاختبار	104.00	604	0.17		
	تطبيق الاختبار	139.70	604	0.23		
	ورقة الاختبار	164.13	604	0.27		
	تصحيح الاختبار	117.57	604	0.19		
	الدرجة الكلية للتقييم	32.63	604	0.05		

يتضح من الجدول السابق ما يلي:

▪ وجود فروق دالة إحصائية (عند مستوى 0.01 أو 0.05) بين متوسطات تقييم جودة المجالات الأربعة (بنية الاختبار، إحصاءات الاختبار، تطبيق الاختبار، ورقة الاختبار) للاختبارات التحصيلية في المرحلة الثانوية بالمملكة ترجع إلى تخصص المعلمين، في حين لا توجد فروق دالة إحصائية بين متوسطات تقييم جودة مجال (تصحيح الاختبار) والدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة ترجع إلى تخصص المعلمين.

جدول (8) نتائج اختبار أقل فرق دال LSD للمقارنات المتعددة لتحديد اتجاه الفروق الدالة إحصائياً عند دراسة اختلاف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف التخصص من وجهة نظر المعلمين

المجال	التخصص	العدد	المتوسط	دراسات أدبية	دراسات علمية	حاسب وتقنية
بنية الاختبار	دراسات أدبية	217	3.67	-		
	دراسات علمية	256	3.77	*0.10	-	

-	*0.13	0.03	3.64	134	حاسب وتقنية	
		-	3.30	217	دراسات أدبية	إحصاءات الاختبار
	-	*0.10	3.40	256	دراسات علمية	
-	*0.14	0.04	3.26	134	حاسب وتقنية	
		-	3.68	217	دراسات أدبية	تطبيق الاختبار
	-	*0.09	3.77	256	دراسات علمية	
-	0.03	*0.12	3.80	134	حاسب وتقنية	
		-	3.83	217	دراسات أدبية	ورقة الاختبار
	-	*0.13	3.70	256	دراسات علمية	
-	0.01	*0.14	3.69	134	حاسب وتقنية	

\*\* تشير إلى أن الفرق بين المتوسطين دال إحصائياً عند مستوى (0.05)

يتضح من الجدول السابق ما يلي:

- أن تقييم جودة (بنية الاختبار) من وجهة نظر المعلمين ذوي التخصص (دراسات علمية) أعلى بدلالة إحصائية من تقييم المعلمين ذوي التخصصات (دراسات أدبية، وحاسب وتقنية).
- أن تقييم جودة (إحصاءات الاختبار) من وجهة نظر المعلمين ذوي التخصص (دراسات علمية) أعلى بدلالة إحصائية من تقييم المعلمين ذوي التخصصات (دراسات أدبية، وحاسب وتقنية).
- أن تقييم جودة (تطبيق الاختبار) من وجهة نظر المعلمين ذوي التخصصات (دراسات علمية، وحاسب وتقنية) أعلى بدلالة إحصائية من تقييم المعلمين ذوي التخصص (دراسات أدبية).
- أن تقييم جودة (ورقة الاختبار) من وجهة نظر المعلمين ذوي التخصص (دراسات أدبية) أعلى بدلالة إحصائية من تقييم المعلمين ذوي التخصصات (دراسات علمية، وحاسب وتقنية).

#### السؤال الرابع:

للإجابة عن السؤال الرابع الذي ينص على: هل يختلف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف سنوات الخبرة في التدريس من وجهة نظر المعلمين؟ تم استخدام:

- تحليل التباين متعدد المتغيرات التابعة Multivariate Analysis of Variance MANOVA
- متبوعاً باختبار أقل فرق دال (Least significant difference) LSD للمقارنات المتعددة، ونتائج هذا السؤال بالجدولين التاليين:

جدول (9) نتائج تحليل التباين متعدد المتغيرات التابعة عند دراسة اختلاف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف سنوات الخبرة في التدريس

مصدر التباين	المجال	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	مستوى الدلالة
التخصص	بنية الاختبار	1.10	3	0.37	1.64	0.18
	إحصاءات الاختبار	2.50	3	0.83	4.85	0.01
	تطبيق الاختبار	12.93	3	4.31	20.23	0.01
	ورقة الاختبار	3.65	3	1.22	4.51	0.01
	تصحيح الاختبار	0.35	3	0.12	0.59	0.62
	الدرجة الكلية للتقييم	1.35	3	0.45	8.63	0.01
الخطأ	بنية الاختبار	134.44	603	0.22		
	إحصاءات الاختبار	103.38	603	0.17		
	تطبيق الاختبار	128.44	603	0.21		
	ورقة الاختبار	162.95	603	0.27		
	تصحيح الاختبار	117.48	603	0.19		
	الدرجة الكلية للتقييم	31.48	603	0.05		

يتضح من الجدول السابق ما يلي:

- وجود فروق دالة إحصائية (عند مستوى 0.01) بين متوسطات تقييم جودة المجالات الثلاثة (إحصاءات الاختبار، تطبيق الاختبار، ورقة الاختبار) والدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة ترجع إلى سنوات الخبرة في التدريس لدى المعلمين. في حين لا توجد فروق دالة إحصائية بين متوسطات تقييم جودة المجالين (بنية الاختبار، تصحيح الاختبار) ترجع إلى سنوات الخبرة في التدريس لدى المعلمين.

جدول (10) نتائج اختبار أقل فرق دال LSD للمقارنات المتعددة لتحديد اتجاه الفروق الدالة إحصائياً عند دراسة اختلاف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف سنوات الخبرة في التدريس من وجهة

#### نظر المعلمين

المجال	سنوات الخبرة	العدد	المتوسط	5 سنوات فأقل	10-6 سنوات	15-11 سنة
إحصاءات الاختبار	5 سنوات فأقل	125	3.21	-	-	-
	10-6 سنوات	246	3.37	*0.16	-	-
	15-11 سنة	135	3.35	*0.14	0.02	-
تطبيق الاختبار	أكثر من 15 سنة	101	3.38	*0.17	0.01	0.03
	5 سنوات فأقل	125	3.62	-	-	-
	10-6 سنوات	246	3.64	0.02	-	-
	15-11 سنة	135	3.96	*0.34	*0.32	-
	أكثر من 15 سنة	101	3.89	*0.27	*0.25	0.07

			3.80	125	5 سنوات فأقل	
		0.15*	3.65	246	10-6 سنوات	ورقة الاختبار
		0.02	3.78	135	15-11 سنة	
		*0.13	3.83	101	أكثر من 15 سنة	
			3.64	125	5 سنوات فأقل	
		0.01	3.65	246	10-6 سنوات	الدرجة الكلية للتقييم
		*0.11	3.75	135	15-11 سنة	
		*0.08	3.73	101	أكثر من 15 سنة	

\*\* تشير إلى أن الفرق بين المتوسطين دال إحصائيًا عند مستوى (0.05)

يتضح من الجدول السابق ما يلي:

- أن تقييم جودة (إحصاءات الاختبار) من وجهة نظر المعلمين منخفضي سنوات الخبرة (5 سنوات فأقل) أقل بدلالة إحصائية من تقييم المعلمين ذوي سنوات الخبرة (6-10 سنوات، 11-15 سنة، أكثر من 15 سنة). أي أنه كلما ارتفع عدد سنوات الخبرة في التدريس ارتفع تقييم المعلمين لجودة إحصاءات الاختبار.
- أن تقييم جودة (تطبيق الاختبار) من وجهة نظر المعلمين ذوي سنوات الخبرة (5 سنوات فأقل، 6-10 سنوات) أقل بدلالة إحصائية من تقييم المعلمين ذوي سنوات الخبرة (11-15 سنة، أكثر من 15 سنة). أي أنه كلما ارتفع عدد سنوات الخبرة في التدريس ارتفع تقييم المعلمين لجودة تطبيق الاختبار.
- أن تقييم جودة (ورقة الاختبار) من وجهة نظر المعلمين ذوي سنوات الخبرة (6-10 سنوات) أقل بدلالة إحصائية من تقييم المعلمين ذوي سنوات الخبرة الأخرى (5 سنوات فأقل، 11-15 سنة، أكثر من 15 سنة).
- أن (الدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة) من وجهة نظر المعلمين ذوي سنوات الخبرة (5 سنوات فأقل، 6-10 سنوات) أقل بدلالة إحصائية من تقييم المعلمين ذوي سنوات الخبرة (11-15 سنة، أكثر من 15 سنة). أي أنه كلما ارتفع عدد سنوات الخبرة في التدريس ارتفعت الدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة.

#### المناقشة والتوصيات:

تعد الاختبارات التحصيلية جزء لا يتجزأ من العملية التعليمية، بل أنها أداء مهمة للوقوف على مدى نجاح العملية التعليمية، لذا هدفت الدراسة إلى الكشف عن مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة في ضوء معايير الاختبار الجيد من وجهة نظر المعلمين، ودراسة اختلاف مستوى جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة باختلاف كل من: النوع (معلم/معلمة)، والتخصص، وسنوات الخبرة في التدريس من وجهة نظر المعلمين، على عينة مكونة من (607) معلم ومعلمة والمرحلة الثانوية بالمملكة العربية السعودية، وطُبق عليهم مقياس جودة الاختبارات التحصيلية، وذلك بعد التحقق من ثباته وصدقه. وتم تحليل

البيانات باستخدام المتوسطات والنسب المئوية واختبار (ت) وتحليل التباين متعدد المتغيرات التابعة. وتوصلت الدراسة إلى مجموعة من النتائج من أهمها:

- أن معلمي ومعلمات المرحلة الثانوية بالمملكة بعينة الدراسة يُقيّمون بوجه عام الاختبارات التحصيلية في المرحلة الثانوية على أنها عالية الجودة؛ حيث تم تقييم المجالات الأربعة (بنية الاختبار، تطبيق الاختبار، ورقة الاختبار، تصحيح الاختبار) على أنها عالية الجودة، بينما تم تقييم مجال (إحصائيات الاختبار) على أنه متوسط الجودة.
- أن متوسط تقييم المعلمات لجودة المجالات الثلاثة (بنية الاختبار، تطبيق الاختبار، ورقة الاختبار) والدرجة الكلية لتقييم جودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة أعلى بدلالة إحصائية من نظائرها لدى المعلمين.
- أن تقييم جودة (بنية الاختبار، إحصاءات الاختبار) من وجهة المعلمين ذوي التخصص (دراسات علمية) أعلى بدلالة إحصائية من تقييم المعلمين ذوي التخصصات (دراسات أدبية، وحاسب وتقنية). في حين أن تقييم جودة (ورقة الاختبار) من وجهة المعلمين ذوي التخصص (دراسات أدبية) أعلى بدلالة إحصائية من تقييم المعلمين ذوي التخصصات (دراسات علمية، وحاسب وتقنية).
- أنه كلما ارتفع عدد سنوات الخبرة في التدريس ارتفع تقييم المعلمين لجودة الاختبارات التحصيلية في المرحلة الثانوية بالمملكة.

وقد أوصت الدراسة بالتالي:

- حث المعلمين على الالتزام بمعايير الاختبار الجيد عند إعداد الاختبارات التحصيلية.
- عقد دورات تدريبية وورش عمل للمعلمين لتدريبهم على خطوات بناء الاختبارات التحصيلية في ضوء معايير الاختبار الجيد.
- عمل ندوات تثقيفية للمعلمين للاهتمام بتطوير الاختبارات التحصيلية ومدى مطابقتها لمعايير الاختبار الجيد.
- تطبيق الدراسة على عينات أخرى بمراحل التعليم المختلفة سواء العام أو الجامعي.

## المراجع العربية:

- إبراهيم، خالد جاسم (2016). *تقويم كفاءة جهاز الإشراف التربوي والاختصاصي في مجال القياس والتقويم التربوي*. مطبعة الإخاء: العراق.
- أبو زيد، حسن سعيد والحسن، ضياء الدين (2017). *تحليل وتقويم الاختبارات التحصيلية المعرفية المستخدمة في تقويم مقررات التربية التقنية ميكانيكا في ضوء مواصفات الاختبار الجيد للأعوام "2013-2017 م". مجلة الدراسات العليا*. مج 11(41). 26-48.
- أبو مديغم، عبد الكريم سالم والعطاونة، ماجد منور (2021). *فاعلية الاختبارات التحصيلية في ضوء معايير بناء الاختبار الجيد من وجهة نظر معلمي المرحلة الثانوية في مدارس بلدة حورة-بئر السبع*. مجلة الفنون والآداب وعلوم الإنسانيات والاجتماع. ع 67. 145-165.
- الجلندانية، رفاء والعميرية، آمنه والراسبية، الريم، وإسماعيل، ميمي (2022). *مستوى رضا أعضاء هيئة التدريس وطالبات جامعة التقنية والعلوم التطبيقية بالبرستاق عن الاختبارات التحصيلية الإلكترونية*. المجلة العربية للقياس والتقويم. ع 5. 119-139.
- الحري، رافدة (2007). *التقويم التربوي الشامل*. دار الفكر للنشر والتوزيع: الأردن.
- الدوسري، راشد (2004). *القياس والتقويم التربوي الحديث*. دار الفكر للنشر والتوزيع.
- الرويلي، مشعل مونس (2016). *مهارات بناء الاختبارات التحصيلية لدى الطلاب المعلمين بالمملكة العربية: دراسة تحليلية تشخيصية من وجهة نظرهم والمتخصصين*. مجلة كلية التربية - جامعة الأزهر. مج 35(170). ج 2. 369-407.
- الزبون، حابس سعد موسى (2017). *مدي كفاءة معلمي المدارس الحكومية الثانوية التابعة لإدارة التربية والتعليم بالطائف في بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد*. مجلة كلية التربية. ع 14. 92-130.
- الزهراني، محمد بن راشد (2009). *تصور مقترح لتطوير أدوات قياس تحصيل الطلاب وفق معايير الجودة الشاملة بوزارة التربية والتعليم*. رسالة دكتوراه غير منشورة، كلية التربية، جامعة أم القرى.
- العابدي، رائد خليل (2006). *الاختبارات المدرسية*. ط 1. مكتبة المجتمع للنشر والتوزيع: الأردن.
- ضيف، حسين والأسود، الزهرة (2021). *كفايات بناء الاختبارات التحصيلية الجيدة لدى أساتذة التعليم الابتدائي دراسة ميدانية على عينة من أساتذة التعليم الابتدائي بولاية الزادي*. مجلة العلوم النفسية والتربوية. مج 2(7). 127-139.
- عبد الواحد، ضياء الدين (2015). *تقويم الطلاب بأقسام المكتبات والمعلومات في ضوء معايير الجودة والاعتماد دراسة تحليلية للاختبارات التحصيلية، المؤتمر الخامس والعشرون: جودة الأداء في المكتبات ومراكز المعلومات والأرشيف*.

- العتوم، محمد صالح (2014). تقييم الاختبارات التحصيلية من إعداد معلم العلوم في محافظة جرش: دراسة تحليلية لنتائج الطلبة للفصل الدراسي الثاني للعام 2013/2012 في مبحث العلوم. مجلة جرش للبحوث والدراسات. 2(15). 361-375.
- عثمان، زليخة عبد العزيز و عبدالله، عبدالرحمن أحمد (2016). تقويم أسئلة امتحانات شهادة التعليم الأساسي بولاية الخرطوم في مادة الفقه والعقيدة للأعوام "2007-2009م". مجلة العلوم التربوية. مج17(2) 80-96.
- عدائكة، دنيا وجخراب، محمد وجعفرور، ربيعة (2018). مستوى معرفة أساتذة التعليم الثانوي لكفايات إعداد الاختبارات التحصيلية الجيدة في ضوء أدوارهم الجديدة. مجلة الباحث في العلوم الإنسانية والاجتماعية. ع33. 1055-1072.
- عزي، إيمان وإسماعيل، راجي وإسعادي، فارس (2018). تحديد الاحتياجات التدريبية في بناء الاختبارات التحصيلية لأساتذة التعليم المتوسط. مجلة الباحث في العلوم الإنسانية والاجتماعية. ع34. 153-168.
- عودة، أحمد (2000). القياس والتقويم في العملية التدريسية. ط3. دار الأمل: الأردن.
- غرايسة، أحمد وعواريب، لخضر (2021). قياس مدى جودة التقويم في المدرسة الجزائرية من خلال بناء الاختبارات التحصيلية (امتحانات التعليم المتوسط نموذجاً). مجلة العلوم النفسية والتربوية. 8(1). 163-172.
- الغزال، محمد عمر (2020). تقويم الاختبارات التحصيلية بمرحلة الدراسات العليا بقسم اللغة العربية وآدابها بكلية الآداب جامعة مصراتة في ضوء معايير الجودة الشاملة. مجلة كلية الآداب. ع15. 31-55.
- فرطاس، خوله (2021). مدى التزام أساتذة التعليم الجامعي بمعايير جودة الاختبار في الاختبارات التحصيلية بقسم العلوم الاجتماعية بجامعة أم البواقي، رسالة ماجستير غير منشورة، كلية العلوم الإنسانية والاجتماعية، جامعة أم البواقي.
- محمد، بلعالية (2021). الاختبارات التحصيلية كأسلوب فعال للتقويم التربوي. مجلة الخلدونية للعلوم الإنسانية والاجتماعية. 13(1). 53-61.
- محمد جاسم، العبيدي العبيدي، محمد جاسم (2011). القياس النفسي والاختبارات. دار الثقافة للنشر والتوزيع: الأردن.
- نصار، سامي (2009). قضايا تربوية في عصر العولمة وما بعد الحداثة. الدار المصرية اللبنانية: القاهرة.
- النوافلة، وليد والعمرى، علي والمهيدات، رزان (2021). درجة ممارسة معلمي العلوم لتحليل نتائج الاختبارات التحصيلية وعلاقته باتجاهاتهم نحوها، مجلة دراسات للعلوم التربوية. مج48(1). 523-538.
- الهويدي، زيد (2004). أساسيات القياس والتقويم التربوي. دار الكتاب الجامعي: العين.



- بدر، محمد أحمد (2017). تقويم الاختبارات المدرسية في ضوء تطوير الاختبارات التحصيلية المدرسية. مطبعة الياسمين: العراق.
- اليعقوبي، حيدر (2013). التقويم والقياس في العلوم التربوية والنفسية. مركز المرتضى للتنمية الاجتماعية: العراق.

#### المراجع العربية مترجمة:

- Abdel Wahed, Diao El Din (2015). Evaluation of students in libraries and information departments in quality and accreditation standards. An analytical study of analytical tests. Twenty-fifth conference: Performance quality in libraries, information centers and archives.
- Abedi, Raed Khalil (2006). school tests. 1st Edition. Community Library for Publishing and Distribution: Jordan.
- Abu Madigham, Abd al-Karim Salem and al-Atawneh, Majed Munawar (2021). The effectiveness of achievement tests in the light of test standards from the point of view of secondary school teachers in the schools of Hura-Beersheba. Journal of Arts, Letters and Humanities. P. 67. 145-165.
- Abu Zayd, Hassan Saeed and Al Hassan, Diao El-Din (2017). Analyzing and evaluating technical knowledge achievement tests in evaluating education courses Mechanics Journal of Graduate Studies. MG 11, p. 41, 26-48.
- Ad'ida, Dunya and Jakharab, Muhammad and Ja'four, Rabia (2018). The level of secondary education teachers' knowledge of the competencies of preparing good achievement tests in the light of their new roles. Journal of Researcher in the Humanities and Social Sciences. P. 33. 1055-1072.
- Al-Atoum, Muhammad Salih (2014). Evaluation of achievement tests prepared by a science teacher in Jerash: an analytical study of conservative results for the second semester of the year 2012/2013 in science research. Jerash Journal for Research and Studies. 2 (15). 375-361.
- Al-Dosari, Rashid (2004). Modern educational measurement and evaluation. Dar Al-Fikr for publication and distribution.
- Al-Ghazal, Muhammad Omar (2020). Evaluation of achievement tests at the postgraduate level in the Department of Arabic Language and Literature at the Faculty of Arts, Misurata University in the light of comprehensive quality standards. College of Arts Journal. P. 15. 31-55.
- Al-Huwaidi, Zaid (2004). Fundamentals of educational measurement and evaluation. University Book House: Al-Ain.
- Al-Jaldaniyah, Refaa and Al-Omairi, Amna and Al-Rasabiyya, Al-Reem, and Ismail, Mimi (2022). The level of satisfaction of faculty members and female students at the University of Technology and Applied Sciences in Rustaq with



- electronic achievement tests. Arab Journal of Measurement and Evaluation. P. 5. 119-139.
- Al-Nawafila, Walid and Al-Omari, Ali and Al-Muhaidat, Razan (2021). The degree of science teachers' practice of analyzing the results of achievement tests and its relationship to their accusations towards them, Journal of Educational Science Studies. Maj. 48. P. 1. 523-538.
  - Al-Ruwaili, Mishaal Mons (2016). Skills of constructing achievement tests among student teachers in the Kingdom of Saudi Arabia: an analytical and diagnostic study from their and specialists' point of view. Journal of the Faculty of Education - Al-Azhar University. MG 35, p. 170, vol. 2, 369-407.
  - Al-Yaqoubi, Haider (2013). Evaluation and measurement in educational and psychological sciences. Al-Mortada Center for Social Development: Iraq.
  - Al-Zahrani, Mohammed bin Rashid (2009). Comprehensive assessment criteria for information and education. Unpublished PhD thesis, College of Education, Umm Al-Qura University.
  - Azi, Eman and Ismail, Rabhi and Essaadi, Fares (2018). Determining the training needs in building achievement tests for middle school teachers. Journal of Researcher in the Humanities and Social Sciences. P. 34. 153-168.
  - Badr, Mohamed Ahmed (2017). Evaluation of school tests in light of the development of school achievement tests. Al-Yasmeen Press: Iraq.
  - Deif, Hussein and Al-Aswad, Al-Zahra (2021). The competencies of building achievement tests for primary education teachers, a field study on a sample of primary education teachers in the state of Zadi. Journal of Psychological and Educational Sciences. P. 7. Mg 2. 127-139.
  - Fartas, Khawla (2021). The quality of university education professors with quality standards in achievement tests at the Department of Social Sciences in Umm El-Bouaghi, unpublished master's thesis, College of Humanities and Social Sciences, Umm El-Bouaghi University.
  - Gharaysa, Ahmed and Aarib, Lakhdar (2021). Measuring the quality of evaluation in the Algerian school through achievement tests (intermediate education exams as a model). Journal of Psychological and Educational Sciences. 8 (1). 163-172.
  - Hariri, Rafida (2007). Comprehensive educational calendar. Dar Al-Fikr for Publishing and Distribution: Jordan.
  - Ibrahim, Khaled Jassim (2016). Evaluation of the efficiency of the educational apparatus and the specialist in educational measurement and evaluation. Brotherhood Press: Iraq.
  - Muhammad Jassim, Al-Obaidi (2011). Psychometric tests. House of Culture for Publishing and Distribution: Jordan.
  - Muhammad, Belalia (2021). Achievement tests as an effective method for educational evaluation. Khaldouniyah Journal of Humanities and Social Sciences. 13 (1). 61-53.

- Nassar, Sami (2009). Educational issues in the era of globalization and postmodernism. The Egyptian Lebanese House: Cairo.
- Odeh, Ahmed (2000). Measurement and evaluation in the teaching process. 3rd Edition, Dar Al-Amal: Jordan.
- Othman, Zulaikha Abdul-Aziz and Abdullah, Abdul-Rahman Ahmed (2016). Evaluation of basic education certificate examination questions in the state of Khartoum in jurisprudence and belief for the years 2007-2009 AD. Journal of Educational Sciences. Vol. 17, p. 2. 80-96.
- The customer, Habis Saad Musa (2017). Secondary schools affiliated with Education and Education in Taif in building tests. College of Education Journal. P. 14. 92-130.