

درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية



وفاء شبيب الدوسري⁽¹⁾

تُسَلَّم البحث في: 2026/1/11

أجيز للنشر في: 2026/3/26

الدوسري، وفاء شبيب. (2026). درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية. *مجلة دراسات الخليج والجزيرة العربية*, 52(201), 243-284.

<https://doi.org/10.34120/jgaps.v52i201.3817>



هذه الدراسة ذات وصول مفتوح نشرت بموجب رخصة المشاع الإبداعي - (CC BY-NC 4.0).

ملخص

الأهداف: هدفت الدراسة إلى الكشف عن درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية، وكذا معرفة اختلاف درجات تلك الممارسة؛ في ضوء المتغيرات الآتية: الخدمة في التدريس، نوع المواد الدراسية، التخصص. **المنهج:** استخدمت الباحثة المنهج الوصفي المسحي، وقد تمثلت أداة الدراسة في مقياس من إعداد الباحثة، طبقته على عينة مكونة من 105 معلمات في المرحلة الثانوية من جميع التخصصات بمحافظة بقيق بالمملكة العربية السعودية؛ تم اختيارهن باستخدام أسلوب الحصر الشامل. **النتائج:** أظهرت نتائج الدراسة أن درجة ممارسة معلمات التعليم العام في المرحلة الثانوية لمهارات تحليل نتائج الاختبارات التحصيلية -بوجه عام- كانت مرتفعة. كما أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية؛ تعزى للمتغيرات الآتية: الخدمة في التدريس، نوع المواد الدراسية، التخصص. **الخاتمة:** قُدمت مجموعة من التوصيات، من أهمها: عقد دورات تدريبية تطبيقية للمعلمات -على اختلاف تخصصاتهن- في مجال التحليل الإحصائي لنتائج الاختبارات التحصيلية؛ لتحسين مهاراتهن في تحليل الاختبارات وتطويرها.

الكلمات المفتاحية: ممارسة المعلمات، الاختبار التحصيلي النهائي، مهارات

تحليل نتائج الاختبار، تحليل فقرات الاختبار

(1) أستاذ مساعد، قسم المناهج وطرق التدريس، كلية التربية، جامعة الملك فيصل، السعودية.

waldossary@kfu.edu.sa

The degree of female general education teachers' practice of skills in analyzing achievement test results

Wafaa Sh. Al-Dossary⁽¹⁾

Abstract

Objectives: This research aimed to examine the degree to which female general education teachers practice skills related to analyzing achievement test results, as well to determine whether this practice differs according to the following variables: years of teaching experience, subject type, and specialization. **Method:** To achieve the research objectives, the researcher employed the descriptive survey approach. The research instrument—a scale developed by the researcher—was administered to a total sample of 105 secondary school female teachers from various disciplines in Buqayq Governorate, Kingdom of Saudi Arabia. **Results:** The findings revealed that, overall, the degree of practice of achievement test result analysis skills among female general secondary education teachers was high. Furthermore, no statistically significant differences were found in the degree of practice attributable to teaching experience, subject type, or specialization. **Conclusion:** In light of these findings, several recommendations were proposed, most notably the organization of practical training courses for female teachers across all specializations in the field of statistical analysis of achievement test results, with the aim of enhancing their skills in analyzing and improving tests.

Keywords: teachers' practice, final achievement test, test result analysis skills, item analysis

Submitted: 11/1/2026

Accepted: 26/3/2026

Al-Dossary, W. Sh. (2026).

The degree of female general education teachers' practice of skills in analyzing achievement test results. *Journal of the Gulf and Arabian Peninsula Studies*, 52(201), 243–284.

<https://doi.org/10.34120/jgaps.v52i201.3817>



This article is an open access article distributed under the Creative Commons Attribution license (CC BY-NC 4.0).

(1) Assistant Professor, Department of Curriculum and Instruction, College of Education, King Faisal University, Saudi Arabia. waldossary@kfu.edu.sa

المقدمة

يُعدّ التقويم التربوي مكوناً رئيساً ودعامة أساسية في العملية التعليمية؛ فهو أحد عناصر المنهج المدرسي، كما أنه الأسلوب العلمي الذي تُعرّف من خلاله جودة النظام التعليمي، ومستوى كفاءة مدخلات العملية التعليمية في تحقيق الأهداف المنشودة، ولا سيما أن عملية التقويم، حينما تتصف بالشمول والتكامل والاستمرارية؛ تقود إلى التغيير والتطوير والتحسين، في جميع جوانب العملية التعليمية.

وعلى الرغم من تعدد أدوات التقويم، تعد الاختبارات التحصيلية من أهم هذه الأدوات، وأكثرها استخداماً وشيوعاً من قبل المعلمين والمعلمات في مدارس التعليم العام؛ فهي تعطي بيانات وإحصاءات دقيقة، تساعد في رسم الخطط التطويرية واتخاذ القرارات لتحسين عمليتي التعليم والتعلم ورفع مستوى دقتها، كما أنها تساعد المعلم في معرفة صعوبات التعلم لدى المتعلمين، وأسبابها، ومستوى تقدمهم نحو تحقيق الأهداف التعليمية، واستثارة دافعيتهم للتعلم، كما أنها تزود المتعلمين بتغذية راجعة عن أدائهم.

ويذكر Sengkaton et al. (2020) أن الاختبار التحصيلي أداة لقياس تحصيل الطلاب، وهو يجري في نهاية البرنامج أو الفصل الدراسي؛ لتحديد مستوى الأهداف التعليمية التي تم تحقيقها، والتأكد من اكتساب الطلاب لنتائج التعلم المقصودة. بينما ترى Fatimah et al. (2020) أن الاختبار التحصيلي: أداة تقويم ذات دور مهم في قياس عملية التعلم والتعليم في المدارس، وذلك من خلال قياس قدرة الطالب، وكذا كفاءة عملية التعلم والتعليم؛ إذ يقدم معلومات للمعلم حول الجوانب التي حقق فيها الطالب نتائج التعلم، وما يعرفه الطلاب ويفهمونه، وما يمكنهم فعله، وما يشعرون به في مراحل مختلفة من عملية التعلم.

وقد اهتمت وزارة التعليم في المملكة العربية السعودية بالاختبارات؛ إذ توجد إدارة تقويم الأداء المعرفي والمهاري في جميع إدارات التعليم، ومن مهامها الاهتمام بالاختبارات بأنواعها. كما تقدم الوزارة عديداً من الدورات التدريبية للمعلمين والمعلمات في مجال الاختبارات؛ لتطوير مهاراتهم وتحسين أدائهم في بناء الاختبارات. إضافة إلى أن الرخصة المهنية تتضمن أن يمتلك المعلمون والمعلمات مهارات بناء الاختبار الجيد، والتحقق من مواصفاته، وكيفية جمع البيانات، وتحليلها، وتفسيرها.

ويأتي الاهتمام بإعداد الاختبارات التحصيلية وفق أسس علمية لكونها تشكّل مهمة أساسية للمعلم؛ ومن ثم، يكون ملماً بخطوات إعدادها، وتتم عملية إعداد الاختبارات التحصيلية بعددٍ من الخطوات؛ يمكن إجمالها فيما يأتي: تحديد الغرض من الاختبار، وتحديد الموضوعات الداخلة في الاختبار وتحليل محتواها المعرفي، وكذلك تحديد نواتج التعلم، وبناء جدول المواصفات، ثم صياغة فقرات (أسئلة) الاختبار بعد اختيار الشكل المناسب (أسئلة مقالية، أو أسئلة موضوعية)، إضافة إلى مراجعة أسئلة الاختبار. وبعد ذلك، تُخرج ورقة الاختبار في صورتها النهائية، ثم يجري الاختبار، وتُقوم إجابات المتعلمين. وأخيراً تُحلّل النتائج وتفسّر (أبو فودة وبني يونس، 2012؛ الجامعة العربية المفتوحة، 2005؛ الزيود وعليان، 2005؛ عباس وآخرون، 2009؛ عودة، 2010؛ منسي، 2007). ويتصف الاختبار الجيد -كما ذكر عطوي (2007) - بالموضوعية؛ إذ إن أداة القياس يجب أن تتسم بالصدق والثبات، والقدرة على التمييز بين الطلاب، وسهولة الاستخدام.

ولأهمية تقويم أداء المعلم في مجال الاختبارات وإعدادها؛ فإن هناك عديداً من الدراسات التي أشارت إلى وجود كثيرٍ من أوجه القصور في معرفة المعلمين للكفايات المتعلقة ببناء الاختبارات التحصيلية وممارستها، ومنها: دراسة الزبون (2013)، ودراسة الفرجات (2015)، ودراسة Yusrizal et al. (2017)، ودراسة Zakariya (2020).

وتعتبر خطوة تحليل نتائج الاختبار من الأنشطة المهمة في تطوير الاختبار للحصول على أسئلة عالية الجودة؛ فهي عملية تهدف إلى تحسين جودة الفقرات؛ من خلال مراجعة الأسئلة، واستبعاد الأسئلة غير الفعالة (Manfaat et al., 2021). ويُعد تحليل فقرات الاختبار مفيداً للمعلمين؛ لتحسين مهاراتهم في بناء اختبارات أكثر صدقاً وثباتاً لقياس تحصيل الطلاب، وفحص المفاهيم الخاطئة، والمحتوى الذي يصعب فهمه من قبل الطلاب ومواطن الضعف المفاهيمي لديهم (علام، 2006؛ عودة، 2010؛ Ahmad, 2019; Sengkaton et al., 2020). (Tsegaye, 2019).

وقد أشار عودة (2010) إلى نوعين من التحليل، يمكن أن يوظفهما المعلم: النوع الأول: التحليل الإجمالي (Wholistic Analysis)؛ وهو التحليل الذي يأخذ العلامة الكلية على الاختبار كوحدةٍ واحدةٍ في التحليل، والنوع الثاني: التحليل الجزئي (Atomistic Analysis)؛ وهو التحليل

الذي يأخذ العلامة على السؤال الواحد أو الفقرة الواحدة كمتغير. ويرتبط التحليل الكلي للاختبار ارتباطاً وثيقاً بنظرية القياس الكلاسيكية، التي تركز على تقدير الخصائص السيكومترية للاختبار ككل؛ مثل معامل الثبات والصدق، بالإضافة إلى المؤشرات الإحصائية العامة التي تعكس مدى دقة الاختبار واتساق نتائجه. أما التحليل الجزئي ل فقرات الاختبار؛ فيُعنى بدراسة خصائص كل فقرة بصورة مستقلة؛ مثل معامل الصعوبة ومعامل التمييز وفاعلية البدائل، وهو ما يمكن تفسيره في ضوء كل من نظرية القياس الكلاسيكية ونظرية الاستجابة للفقرة؛ إذ تسهم هاتان النظريتان في تفسير أداء الفقرات وعلاقتها بقدرات المفحوصين (علام، 2006؛ عودة، 2010).

ويتضمن التحليل الإجمالي عدة معالجات إحصائية؛ تعطي وصفاً بيانياً؛ مثل: ترتيب الدرجات وتبويبها وعرض الدرجات بصورة جدول أو توزيع تكراري، ثم تحويله إلى مدرج تكراري، أو مضلع تكراري، أو منحني تكراري، ثم إعطاء وصف كمي لهذا التوزيع؛ من خلال مقاييس النزعة المركزية ومقاييس التشتت. وهذه التحليلات الإحصائية تتعلق بمتغير واحد؛ وهو الدرجات على اختبار واحد.

إن ترتيب الدرجات تصاعدياً أو تنازلياً، يفيد المعلم في معرفة موقع كل طالب مقارنة بغيره. كما يوضح أعلى درجة وأقلها، والدرجات التي يشترك فيها الطلاب في الرتبة. وقد يعكس هذا الترتيب صعوبة الاختبار، ومستوى الطلاب. كما قد تُعرض الدرجات الخام في صورة جدول تكراري يترجم إلى تمثيل بياني (أبو لبدة، 2008؛ العدل، 2017؛ عطية، 2010؛ عودة، 2010؛ فليفل وحمدان 2009؛ مراد، 2000).

ويذكر عودة (2010) أن الصورة الكاملة للاختبار تكون أوضح؛ من خلال استخدام الوصف الكمي للتوزيع، إلى جانب الوصف البياني. ويقصد بالوصف الكمي: تلخيص عدد كبير من الدرجات بقيمة واحدة، أو بعدد قليل من القيم؛ بحيث تحسب باستخدام مقاييس النزعة المركزية (المتوسط، الوسيط، المنوال)، التي توضح مستوى أداء (تحصيل) الطلاب، ومقاييس التشتت (المدى، الانحراف المتوسط، التباين، الانحراف المعياري)، التي تبين مدى تجانس الطلاب في أدائهم الاختبار.

في المقابل، يتناول التحليل التفصيلي فقرات الاختبار؛ من حيث: معاملات الصعوبة، والتميز لكل فقرة، وفاعلية البدائل في حال الأسئلة ذات الإجابات المنتقاة. فمعامل الصعوبة: هو مؤشر يحدد مدى قدرة المتعلمين على الإجابة الصحيحة عن فقرة معينة في الاختبار، وكلما انخفضت قيمة معامل الصعوبة، زادت صعوبة الإجابة عن السؤال لدى الطلاب، والعكس صحيح (Aisyah, 2019 ؛ Fatimah et al., 2020).

بينما يوضح معامل التمييز قدرة المفردة (السؤال أو الفقرة) على التمييز بين أداء الطلاب ذوي التحصيل المرتفع والطلاب ذوي التحصيل المنخفض. أما فاعلية البدائل أو المموهات؛ فهي إجابات غير صحيحة؛ ترد كبدايل للإجابة عن سؤال اختبار من نوع الاختيار من متعدد (العمر، 2007)، والبديل الخطأ الذي يجتذب عدداً من المفحوصين يعد بديلاً فعالاً أو جذاباً؛ أما البديل الخطأ الذي لا يختاره أحد من المفحوصين، فإنه يعد بديلاً غير فعال، أو غير جذاب.

وبناءً على ذلك؛ يمكن القول: إن العلاقة بين التحليل الجزئي والتحليل الكلي للاختبارات التحصيلية علاقة تكاملية؛ إذ يُستخدم كل منهما لفحص جودة الاختبار من زاوية مختلفة لكنهما مترابطة. وتظهر أهمية هذه العلاقة في أن نتائج التحليل الجزئي تسهم في تحسين جودة الاختبار الكلي؛ إذ تؤدي المفردات الجيدة إلى رفع ثبات الاختبار وصدقه، بينما قد تؤدي المفردات الضعيفة إلى خفض كفاءته. لذلك فإن الجمع بين التحليلين يُعد ضرورياً لبناء الاختبارات التحصيلية وتقويمها بما يحقق دقة القياس التربوي.

وفي ضوء ما سبق؛ تبرز أهمية تفعيل الاستفادة من التقنية وجهاز الحاسب الآلي؛ لإعداد جدول المواصفات للاختبار التحصيلي، وتحليل نتائجه. ومن تلك البرامج، برنامج الإكسل (Excel)، وهو أحد برامج مايكروسوفت أوفيس (Microsoft Office)، ويعد من البرامج المهمة التي يستخدمها المعلم في إعداد الجداول، وكشوف متابعة الطلاب، وكذلك الحضور والغياب، إضافة إلى حساب مجموع الدرجات التي حصل عليها الطلاب. كما يمكن من خلاله، استخدام مجموعة واسعة من الصيغ والدوال لإجراء عمليات حسابية معقدة أو بسيطة. وكذلك، إعداد التقارير والإحصائيات مدعمة بالرسوم البيانية، التي توضح المعدلات، ومستويات الطلاب.

ويعد برنامج الحزم الإحصائية للعلوم الاجتماعية (SPSS) من أشهر البرامج الإحصائية التي تُستخدم في تحليل جميع أنواع البيانات؛ فهو يتيح إدخال البيانات التي يتم جمعها

من خلال الاستبانات أو المقابلات أو الملاحظات، ثم تحليلها باستخدام أدوات وقوائم مخصصة؛ تسهل استخراج النتائج بدقة. وكذلك، الوصول إلى مجموعة متنوعة من الأدوات الإحصائية؛ مثل: التحليل الوصفي، والاختبارات الاستدلالية، والتحليل الترابطي، وغيرها. كما يوفر البرنامج، أدوات لعرض النتائج بشكل مرئي؛ مثل: الجداول، والرسوم البيانية؛ مما يساعد في تفسير البيانات، وتقديمها بشكل مبسط وسهل الفهم.

بينما يستخدم برنامج Iteman (Item Analysis) في تحليل فقرات الاختبار، وتقدير أداء الاختبار ككل؛ فهو يتناول تحليلاً تفصيلياً لفقرات الاختبار؛ من حيث: معاملات الصعوبة، والتمييز لكل فقرة، وفاعلية البدائل في حالة الأسئلة ذات الإجابات المنتقاة. وكذلك، يوفر إحصاءات وصفية؛ مثل: المتوسط الحسابي، والانحراف المعياري، ومعامل الثبات، والخطأ المعياري للقياس، وغيرها. كما أنه ينتج تقارير تقنية مفصلة في صيغة مايكروسوفت وورد (Microsoft Word)؛ تتضمن رسومات بيانية مفصلة، إضافة إلى الجداول (Thompson & Guyer, 2010).

بالإضافة إلى ذلك يمثل ماستري (Mastery) برنامج تصحيح آلي للاختبارات مدعوم باللغة العربية. ويتميز البرنامج بقدرته على التحليلين الجزئي والكلّي للاختبارات المدرسية؛ فهو يسمح بتصميم الاختبارات المقالية والموضوعية، كما أنه منصة سعودية؛ مما يعني أنه يتيح إمكانية تصدير الدرجات مباشرة إلى نظام نور بسهولة. ويمكن من خلاله أيضاً، تصحيح أوراق الطلاب من خلال: مسح صور أوراق الإجابة، ورفعها على المنصة، واستخراج تقارير مفصلة بصيغ متنوعة؛ مثل: PDF و Excel (Professional Development Company, 2024). وهناك عديد من التطبيقات والبرامج الإلكترونية لتحليل نتائج الاختبارات التحصيلية، يمكن للمعلمين اختيار ما يناسبهم منها؛ بما يساهم في جودة الاختبار، وتحسين بناء فقراته.

ونتيجةً لقلّة الدراسات التي تناولت درجة ممارسة المعلمات لمهارات تحليل نتائج الاختبارات التحصيلية بصورة مباشرة ومفصلة للتحليل الكلّي والتفصيلي معاً؛ فقد لجأت الباحثة إلى عرض لدراسات مقارنة -نوعاً ما- تناولت تحليل بعض الباحثين لأسئلة اختبارات من إعداد المعلمين، وتم تناول التحليل التفصيلي في هذه الدراسات. وكذلك، لجأت الباحثة - أيضاً- إلى دراسات اهتمت بكفايات بناء الاختبارات التحصيلية، وكفايات القياس والتقويم

لدى المعلمين والمعلمات، وكذا الاحتياجات التدريبية في بناء الاختبارات التحصيلية. وسيكون التركيز في هذه الدراسات، على خطوة التحليل.

أجرى أبو جراد (2011) دراسة هدفت إلى معرفة مدى التزام المعلمين في فلسطين بتحليل نتائج الاختبارات التحصيلية، وعلاقته باتجاهاتهم نحوها، وقد اتبعت الدراسة المنهج الوصفي، وكان عدد أفراد العينة 249 معلماً ومعلمة من المرحلة الأساسية الدنيا في غزة، وتكونت أدوات الدراسة من مقياس الاتجاهات، ومقياس الممارسات الإحصائية. ومن أهم النتائج: وجود ضعف واضح في ممارسة المعلمين لتحليل نتائج الاختبارات التحصيلية، وعدم وجود فروق دالة إحصائية في درجة التزام المعلمين بتحليل نتائج الاختبارات التحصيلية، تعود لمتغير الجنس أو الخبرة.

كما أجرى الزبون (2013) دراسة هدفت إلى استقصاء الحاجات التدريبية لمعلمي المدارس الحكومية الثانوية بالطائف في مجال بناء الاختبارات التحصيلية، وأهم المعوقات التي تحول دون بنائها وفق معايير الاختبار الجيد. اتبعت الدراسة المنهج الوصفي، وكان عدد أفراد العينة 280 معلماً، اختبروا بطريقة العينة العشوائية العنقودية. تكونت أداة الدراسة من اختبار تحصيلي، واستبانة لمسح الحاجات التدريبية للمعلمين وكلاهما في مجالي الاختبارات والمقاييس. ومن أهم النتائج: أن درجة الحاجات التدريبية للمعلمين في مجال بناء الاختبارات التحصيلية كانت متوسطة، وأن هناك حاجة في مجال عقد دورات تدريبية، فيما يخص بناء الاختبارات التحصيلية وبنوك الأسئلة وتحليل نتائجها إحصائياً، وأن أهم المعوقات تتعلق بعدم القدرة على بناء الاختبارات وتحليل نتائجها وتفسيرها.

وهدف دراسة Yusrizal et al. (2017) إلى تحديد مستوى معرفة معلمي الفيزياء وفهمهم لتقويم التعلم، بالإضافة إلى وصف مهاراتهم في بناء فقرات الاختبار وتحليلها. اتبعت الدراسة المنهج الوصفي المسحي، وجمعت البيانات من خلال استبانة. تكونت عينة الدراسة من 32 معلماً في المرحلة الثانوية بمدينة باندا آتشيه في إندونيسيا، اختبروا بطريقة عشوائية من مجتمع الدراسة. ومن أهم النتائج: أن معلمي الفيزياء في خمس مدارس فقط؛ هم من أظهروا مستوى "مرتفعاً جداً" في المعرفة والفهم بتقويم التعلم، وكان عدد المدارس 16 مدرسة. أما مهارات بناء فقرات الاختبار وتحليله لدى معلمي الفيزياء في المرحلة الثانوية، فقد تبين أنها لا تزال دون المستوى المطلوب.

أمّا دراسة عدائكة وآخرين (2018): فقد هدفت إلى تعرّف مستوى معرفة أساتذة التعليم الثانوي في الجزائر لكفايات إعداد الاختبارات التحصيلية الجيدة في ضوء أدوارهم الجديدة. اتبعت الدراسة المنهج الوصفي، وكان عدد أفراد العينة 232 معلماً ومعلمة من ثانويات ولاية الوادي وسط. تكونت أداة الدراسة من مقياس الكفايات. ومن أهم النتائج: أن مستوى معرفة أساتذة التعليم الثانوي لكفايات التحليل والتطبيق في إعداد الاختبارات التحصيلية الجيدة -في ضوء أدوارهم الجديدة- كانت منخفضة.

وسعت دراسة المسعودي والياسر (2018) إلى تحديد كفايات القياس والتقويم لدى معلمة التعليم العام في المراحل الثلاث بتيوك في المملكة العربية السعودية، وفقّ المعايير العالمية. في ضوء بعض المتغيرات. اتبعت الدراسة المنهج الوصفي، وقد بنيت أداتان: استبانة للكشف عن مدى إدراك المعلمات لكفايات القياس والتقويم، وقد طُبِّقَت على 357 معلمة، وبطاقة ملاحظة؛ للكشف عن ممارسة المعلمات لكفايات القياس، وقد طبقت على 60 معلمة، طبقت عليهن أداة الاستبانة مسبقاً. ومن أهم النتائج: أن هناك درجة توافر مرتفعة لمدى إدراك المعلمة لكل معيار من معايير المقياس بشكل عام. أما بطاقة الملاحظة؛ فكانت درجة التوافر متوسطة، ما عدا "كفاية تفسير نتائج عملية القياس والتقويم"؛ إذ احتوت على بعض الممارسات لتحليل الاختبار بدرجة قليلة. وكذلك عدم وجود فروق دالة إحصائية في المتوسطات لدرجة ممارسة المعلمة لكفايات القياس والتقويم، تعزى إلى التخصص؛ بينما تشير إلى وجود فروق تعود إلى كلّ من الخبرة وعدد الدورات.

كما أجرى النوافلة وآخرون (2021) دراسة هدفت إلى تعرّف درجة ممارسة معلمي العلوم في الأردن لتحليل نتائج الاختبارات التحصيلية، وعلاقته باتجاهاتهم نحو الاختبارات. اتبعت الدراسة المنهج الوصفي، وكان عدد أفراد العينة 221 معلماً ومعلمة؛ اختبروا بطريقة عشوائية، في المدارس الأساسية العليا والثانوية في إربد. تكونت أدوات الدراسة من: مقياس الممارسة، واستبانة الاتجاهات. ومن أهم النتائج: أن درجة الممارسة كانت منخفضة، ولم توجد فروق ذات دلالة إحصائية في درجة الممارسة، تعزى إلى أثر كلّ من الجنس أو سنوات الخبرة، ولكن وجدت فروق في المؤهل التربوي، لصالح المعلمين الذين يحملون مؤهلاً تربوياً.

وجاءت دراسة Kissi et al. (2023) للكشف عن العلاقة بين كفاءة المعلمين في بناء اختبارات الاختيار من متعدد وجودة فقرات تلك الاختبارات، لدى معلمي المرحلة الثانوية

العليا في غانا. وقد اتبعت الدراسة المنهج الوصفي والمنهج النوعي، وكان عدد أفراد العينة 157 معلماً، من أربع مدارس ثانوية في منطقة كواهيو-ساوث. تكونت أدوات الدراسة من استبانة؛ بهدف قياس كفاءات المعلمين في بناء فقرات الاختيار من متعدد، مكونة من ثلاثة أبعاد؛ تمثل كفاءة المعلم في بناء الاختبارات. بعدها، قام الباحثون بتحليل اختبارات المعلمين، ومن أهم النتائج: إشارة المعلمين إلى أنهم يشعرون بكفاءة أعلى في ضمان صدق المحتوى، تلبها كفاءتهم في تجميع فقرات الاختبار، ثم كفاءتهم في التعامل مع بدائل الفقرات. وكشفت الدراسة عن وجود مشكلات كبيرة في النسخ الفعلية من اختبارات الاختيار من متعدد التي قام المعلمون ببنائها؛ إذ أظهر تحليل عينة من بنود الاختبار الفعلية عكس ذلك؛ مما يشير إلى وجود فجوة بين الكفاءة الذاتية المدركة وجودة التطبيق الفعلي.

وهدفـت دراسة Ogunsanmi et al. (2025) إلى تحديد صلاحية محتوى اختبار التحصيل الذي يستخدمه المعلمون، ومهارات بناء الاختبار لدى المعلمين، في مدارس ولاية أوسون الثانوية بنيجيريا. اتبعت الدراسة المنهج الوصفي، وتكونت أدوات الدراسة من اختبار مهارات البناء، وبطاقة الملاحظة؛ لجمع البيانات. وكان عدد أفراد العينة 120 معلماً، اختبروا بطريقة قصدية من ست مناطق؛ بناءً على خمس مواد أساسية مقدمة في كل مدرسة، واختيرت 12 مدرسة باستخدام إجراء أخذ العينات العشوائية البسيطة. ومن أهم النتائج: أن مستوى مهارات بناء الاختبارات كان مرتفعاً، وأن متوسط قيمة الكفاءة للمعلمين كان مناسباً بما يكفي لضمان كفاءة مناسبة لاختبارات التحصيل. كما أن مهارات بناء الاختبارات لدى معلمي المدارس الثانوية كانت كافية تماماً لتوفير كفاءة كافية لاختبارات التحصيل.

مما سبق عرضه تكشف أغلب الدراسات السابقة على اختلاف بيئات تطبيقها، أن هناك ضعفاً ملحوظاً لدى المعلمين في ممارسة تحليل الاختبارات التحصيلية، كما أن مهارات بناء فقرات الاختبار وتحليله لا تزال دون المستوى المطلوب، وأنه يمكن مراجعة عديد من فقرات الاختبار لتجويدها واستخدامها في الاختبارات المستقبلية، ما عدا دراسة Ogunsanmi et al. (2025)؛ التي أظهرت أن مهارات بناء الاختبارات لدى معلمي المدارس الثانوية، كانت كافية تماماً لتوفير كفاءة كافية لاختبارات التحصيل.

وتتفق هذه الدراسة مع الدراسات السابقة، في استخدام المنهج الوصفي، وفي المرحلة الثانوية، ما عدا دراسة أبو جراد (2011). كما تتفق مع دراسة كل من: أبو جراد (2011)،

وعداثة وآخري (2018)، والنوافلة وآخري (2021)؛ في استخدامها أداة الدراسة "المقياس"، وفي جزء من العينة؛ وهي المعلمات؛ بالإضافة إلى دراسة المسعودي والجاسر (2018).

تختلف هذه الدراسة عن الدراسات السابقة؛ في كون العينة تمثل مجتمع الدراسة. كما أنها تختلف عن الدراسات السابقة في تناولها درجة ممارسة المعلمات لمهارات تحليل نتائج الاختبارات التحصيلية بصورة مباشرة ومفصلة للتحليل الكلي والجزئي؛ وهي بذلك تتفق -نوفاً ما- مع دراسة أبو جراد (2011)، ودراسة النوافلة وآخري (2021)، ولكن تختلف عنهما في صياغة بعض فقرات المقياس، وفي بعض متغيرات الدراسة؛ ومن ثم، تتميز هذه الدراسة بتقديم إطار نظري متكامل لمهارات تحليل نتائج الاختبارات التحصيلية من خلال التحليلين الجزئي والكلي وإبراز العلاقات بينهما، وربطهما بالأطر النظرية في القياس والتقويم التربوي، لا سيما النظرية الكلاسيكية للاختبار ونظرية الاستجابة للفقرة؛ بما يسهم في تقديم معالجة أكثر تكاملاً لتحليل الاختبارات وتحسين جودة بنائها، مع توظيف برامج تقنية داعمة لتحليل الاختبارات بدقة؛ بحيث تسهم في تطوير مهارات المعلمات وتعزيز فاعلية التقييم المدرسي.

وبناءً على ما سبق، وكذلك لأهمية تحليل نتائج الاختبارات التحصيلية التي يعدها المعلمون، وخاصة في المرحلة الثانوية، لا سيما أن المتعلمين في هذه المرحلة يواجهون اختبارات القدرات والتحصيل من قبل المركز الوطني للقياس لتأهيلهم للمرحلة الجامعية؛ فإن ذلك يتطلب إجراء مزيدٍ من الدراسات حول مهارات تحليل نتائج الاختبارات المدرسية، وإعداد اختباراتٍ تستوفي متطلبات خصائص الاختبار الجيد.

مشكلة الدراسة

إن تحليل نتائج الاختبارات التي يعدّها المعلمون في نهاية كل فصل دراسي (التقويم الختامي) وتفسيرها يُعد من المهارات الأساسية التي ينبغي أن يكتسبها المعلم؛ ذلك أن عدم الاهتمام بها من قبل المعلم ينعكس على الخصائص السيكومترية لتحليل نتائج الاختبار؛ وهو ما أشار إليه Manfaat et al. (2021) بأنه يتم الحصول على نتائج التعلم باستخدام اختبار لا يصف إنجازات الطالب في التعلم؛ وهو ما يؤدي إلى عدم توافق نتائج التعلم مع قدرات الطلاب؛ ومن ثم، يترتب عليه اتخاذ قرارات غير دقيقة، وذات آثار سلبية على الطلاب.

لقد توصلت نتائج الدراسات التي أجريت حول تحليل نتائج الاختبارات التحصيلية إلى أنها لا تزال دون المستوى المطلوب؛ مثل: نتائج دراسة Yusrizal et al. (2017)، وبعضها غير قابلة للاستخدام أداة لتقييم أداء الطلاب؛ مثل نتائج دراسة Sengkaton et al. (2020)، في حين أظهرت نتائج دراسة أبو جراد (2011) ضعفاً واضحاً في التزام المعلمين بتحليل نتائج الاختبارات التحصيلية.

كما أجرت الباحثة دراسة استطلاعية على معلمات التعليم العام في المرحلة الثانوية، وعددهن 42 معلمة؛ لحصر الممارسات المستخدمة من قبل المعلمات بعد الانتهاء من التصحيح للاختبار. وأظهرت نتائج الدراسة، أن هناك بعض الممارسات للتحليل الكلي من قبل المعلمات، وأغلبها تتمثل في الرسم البياني لتصنيف التقديرات، وحساب متوسط الدرجات، وأعلى درجة وأدناها، ولا توجد ممارسة لتحليل فقرات الاختبار؛ من حيث حساب معامل الصعوبة، ومعامل التمييز، وفعالية المموهات.

هذا، إضافةً إلى توصيات بعض الدراسات؛ مثل دراسة Yusrizal et al. (2017)، ودراسة Zakariya (2020)، فقد أوصتا بضرورة تنظيم ورش عمل ودورات تدريبية في أثناء الخدمة للمعلمين، في مجالات بناء الاختبارات وتصحيحها. وكذا دراسة أبو جراد (2011)، التي أكدت إعادة النظر في الدورات التدريبية الخاصة بإعداد الاختبارات التحصيلية؛ لتكون أكثر عملية وتركيزاً على الجوانب التطبيقية في تحليل النتائج.

واستناداً إلى ما سبق؛ فإن مشكلة الدراسة تحددت في: تدني ممارسة المعلمات لمهارات تحليل نتائج الاختبارات التحصيلية. وللتصدي لهذه المشكلة، فإن هذه الدراسة تحاول الكشف عن درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية.

أسئلة الدراسة

سعت هذه الدراسة إلى الإجابة عن الأسئلة الآتية:

- 1 - ما مهارات تحليل نتائج الاختبارات التحصيلية الواجب توافرها لدى معلمات التعليم العام؟
- 2 - ما درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؟
- 3 - هل تختلف درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ باختلاف كل من: الخدمة في التدريس، نوع المواد الدراسية، التخصص؟

فروض الدراسة

- 1 - توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$) بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير الخدمة في التدريس.
- 2 - توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$) بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير نوع المواد الدراسية.
- 3 - توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$) بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير التخصص.

أهداف الدراسة

هدفت هذه الدراسة إلى:

- تقديم قائمة بأهم مهارات تحليل نتائج الاختبارات التحصيلية؛ الواجب توافرها لدى معلمات التعليم العام.
- الكشف عن درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية.
- معرفة اختلاف درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ باختلاف كل من: الخدمة في التدريس، نوع المواد الدراسية، التخصص.

أهمية الدراسة

تمثلت أهمية الدراسة في أنها:

- تفيد المعلمين في تحسين ممارساتهم لتحليل نتائج الاختبارات التحصيلية وتطويرها وتقييمها. ومن ثم؛ تساهم في تحسين جودة تعلم طلابهم، وبناء اختبارات تحقق خصائص الاختبار الجيد.
- تفيد القائمين على برامج الجودة في تحسين نواتج التعلم؛ من خلال عقد دورات تدريبية تطبيقية وورش عمل خاصة بتحليل الاختبارات التحصيلية.

- قد تساعد متخذي القرار وراسي السياسات التعليمية في التخطيط لبرامج إعداد المعلمين، والاهتمام بالمهارات التي ينبغي أن يمتلكها الطالب المعلم في مجال القياس والتقويم.
- تفتح المجال أمام بحوث ودراسات أخرى لتناول جوانب مرتبطة بتحليل الاختبارات التحصيلية.
- تُقدم أداة بحثية يمكن الاستفادة منها في الدراسات البحثية المستقبلية المشابهة.

مصطلحات الدراسة

الممارسة (practice)

عرّف صبري (2002) الممارسة بأنها: "نوع من الخبرة المنظمة نسبياً، وهي التي تشير إلى تكرار حدوث نفس الاستجابات الظاهرة، أو ما يشهدها في مواقف بيئية منظمة نسبياً" (ص.525). وتُعرّف الممارسة إجرائياً، بأنها: "أداء معلمة المرحلة الثانوية واستخدامها لمجموعة من المعالجات الإحصائية؛ مثل: حساب المتوسط الحسابي، ومعامل الصعوبة، ومعامل التمييز، ... بشكل متكرر ومنظم؛ وذلك لتحليل نتائج الاختبار التحصيلي الذي تعدّه للطالبات في نهاية كل فصل دراسي، في المقررات التي تقوم بتدريسها".

الاختبار التحصيلي (Achievement Test)

عرّف عودة (2010) الاختبار التحصيلي بأنه: "طريقة منظمة لتحديد مستوى تحصيل الطالب لمعلومات ومهارات في مادة دراسية كان قد تعلمها مسبقاً بصفة رسمية، من خلال إجاباته عن عينة من الأسئلة (الفقرات)، التي تمثل محتوى المادة الدراسية" (ص.63). وتُعرّف الاختبارات التحصيلية إجرائياً، بأنها: "الاختبارات الختامية الفصلية التي تجريها معلمات المرحلة الثانوية في نهاية كل فصل دراسي؛ وذلك لتحديد مستوى تحصيل الطالبات لمعلومات ومهارات قد تعلمتها مسبقاً، في مادة دراسية معينة بشكل رسمي".

تحليل نتائج الاختبارات التحصيلية (Analysis of Achievement Test Results)

عرّف النوافلة وآخرون (2021) تحليل نتائج الاختبارات التحصيلية، بأنه: مجموعة من الإجراءات والممارسات الإحصائية التي يستخدمها المعلمون على مستوى الاختبار ككل، وعلى مستوى كل فقرة من فقرات الاختبار، من أجل دراسة الخصائص

السيكومترية للاختبار، ومعرفة الأهداف التي تحققت والأهداف التي لم تتحقق، ونقاط الضعف والقوة لدى الطلبة، وترتيب الطلبة بحسب أدائهم، ومدى مناسبة الاختبار، والمشكلات التي يواجهها الطلبة (ص. 528).

وسوف تأخذ الباحثة بهذا التعريف، مع إجراء بعض التعديلات؛ ومن ثم، يُمكن تعريف ممارسة مهارات تحليل نتائج الاختبارات التحصيلية إجرائياً، بأنها: أداء معلمة المرحلة الثانوية وقدرتها على استخدام مجموعة من الإجراءات والمعالجات الإحصائية على مستوى الاختبار ككل، وعلى مستوى كل فقرة من فقرات الاختبار بسهولة وسرعة ودقة وبصورة منظمة؛ من أجل دراسة الخصائص السيكومترية للاختبار، ومعرفة نواتج التعلم التي تحققت، ونقاط الضعف والقوة لدى الطالبات، وترتيب الطالبات بحسب أدائهن، وتحديد مستوى صعوبة فقرات الاختبار وقدرتها على التمييز، وفعالية البدائل في أسئلة الاختبار من متعدد، بما يمكنها من اتخاذ قرارات تربوية تسهم في تحسين نواتج التعلم؛ وذلك في الاختبار التحصيلي الذي تعده للطالبات في نهاية كل فصل دراسي، في المقررات التي تقوم بتدريسها.

وهي تُقاس في هذه الدراسة، بالدرجة التي تحصل عليها المعلمة على مقياس ممارستها لمهارات تحليل نتائج الاختبارات التحصيلية التي تعدها لطالباتها.

حدود الدراسة

حُدِّدت أطر هذه الدراسة بالحدود الآتية:

- الحدود الموضوعية: تتمثل في معرفة درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية. واقتصرت هذه الدراسة، على أبرز المعالجات الإحصائية التي يمكن أن تستخدمها المعلمة في تحليل نتائج الاختبار التحصيلي.
- الحدود الزمانية: نُفذت هذه الدراسة، في نهاية الفصل الدراسي الثاني، من العام الدراسي 1446هـ / 2025.
- الحدود المكانية: طُبِّقت هذه الدراسة، في مدارس التعليم العام بالمرحلة الثانوية، بمحافظة بقيق في المملكة العربية السعودية.
- الحدود البشرية: طُبِّقت هذه الدراسة، على جميع معلمات المرحلة الثانوية بمحافظة بقيق، في جميع التخصصات، وكان إجمالي عددهن 105 معلمات.

المنهج

استخدم المنهج الوصفي المسحي لإجراء هذه الدراسة؛ من خلال جمع البيانات بواسطة أداة الدراسة (المقياس).

مجتمع الدراسة وعينتها

تكون مجتمع الدراسة من جميع معلمات المرحلة الثانوية، في كل التخصصات، في محافظة بقيق، وعددهن 105 معلمات؛ بحسب آخر إحصائية للعام الدراسي 1446هـ - 2025. أما عينة الدراسة؛ فاختيرت باستخدام أسلوب الحصر الشامل؛ أي إن عينة الدراسة كانت هي مجتمع الدراسة، ومعروف أن استخدام أسلوب الحصر الشامل يحقق عدة مزايا، من أهمها: الحصول على دقة عالية ومباشرة في النتائج؛ مما يزيد من صدق نتائج الدراسة، وإمكانية الاستفادة منها على أرض الواقع، وهو ما أشار إليه العساف (2006). ويوضح جدول 1 توزيع أفراد مجتمع الدراسة وعينتها، بحسب المتغيرات الديموغرافية.

جدول 1

توزيع أفراد مجتمع الدراسة وعينتها بحسب المتغيرات الديموغرافية (الخدمة في التدريس، ونوع المواد الدراسية، والتخصص) (ن = 105)

المتغير	المجموعات الفرعية للمتغير	التكرار	%
الخدمة في التدريس	أقل من خمس سنوات	36	34.29
	من 5 سنوات إلى أقل من 10 سنوات	23	21.90
	من 10 سنوات فأكثر	46	43.81
نوع المواد الدراسية	علمية	56	53.33
	أدبية	39	37.14
	متنوعة (علمية وأدبية)	10	9.52

تابع/ جدول 1

توزيع أفراد مجتمع الدراسة وعينتها بحسب المتغيرات الديموغرافية (الخدمة في التدريس، ونوع المواد الدراسية، والتخصص) (ن=105)

المتغير	المجموعات الفرعية للمتغير	النكرار	%
التخصص	اللغات (اللغة العربية، اللغة الإنجليزية)	27	25.71
	العلوم الطبيعية (الفيزياء، الكيمياء، الأحياء)	33	31.43
	الرياضيات والمهارات الرقمية (الرياضيات، الحاسب الآلي، التقنية الرقمية)	23	21.90
	الدراسات الإسلامية والعلوم الاجتماعية (الدراسات الإسلامية، التربية الأسرية، علم النفس، علم الاجتماع، المكتبات، التاريخ، تاريخ فرع إدارة، الإدارة، الاقتصاد المنزلي)	22	20.95
طريقة تحليل نتائج الاختبارات التحصيلية	برامج إلكترونية	99	94.29
	يدوياً	6	5.71

يتضح من جدول 1:

- أن أعلى نسبة من عينة الدراسة من معلمات المرحلة الثانوية، كانت سنوات الخدمة في التدريس لديهن من 10 سنوات فأكثر؛ إذ بلغت 43.81%. تلتها نسبة المعلمات ذوات سنوات الخدمة في التدريس أقل من خمس سنوات، وقدرها 34.29%. وأخيراً، نسبة المعلمات ذوات سنوات الخدمة في التدريس من 5 سنوات إلى أقل من 10 سنوات، وقدرها 21.90% من إجمالي حجم العينة.
- أن أكثر من نصف عينة الدراسة من معلمات المرحلة الثانوية، كنَّ يقيمون بتدريس مواد علمية؛ إذ بلغت نسبتهن 53.33%. تلتها نسبة المعلمات اللاتي يدرسن مواد أدبية، وقدرها 37.14%. وأخيراً، نسبة المعلمات اللاتي يدرسن مواد متنوعة (علمية وأدبية)، وقدرها 9.52% من إجمالي حجم العينة.
- أن هناك تقارباً بين نسب المعلمات ذوات التخصصات المختلفة في عينة الدراسة، وإن كانت أعلى نسبة منهن من ذوات تخصص "العلوم الطبيعية"؛ إذ بلغت 31.43%. تلتها نسبة المعلمات ذوات تخصص "اللغات"، وقدرها 25.71%، ثم نسبة المعلمات ذوات

تخصص "الرياضيات، والمهارات الرقمية"، وقدرها 21.90%. وأخيراً، نسبة المعلمات ذوات تخصص "الدراسات الإسلامية، والعلوم الاجتماعية"؛ وقدرها 20.95%، من إجمالي حجم العينة.

– أن معظم عينة الدراسة من معلمات المرحلة الثانوية، كُنَّ يَقُمن بتحليل نتائج الاختبارات التحصيلية باستخدام برامج إلكترونية؛ إذ بلغت نسبتهن 94.29%، في حين بلغت نسبة المعلمات اللاتي كُنَّ يَقُمن بتحليل نتائج الاختبارات التحصيلية بصورة يدوية 5.71%، من إجمالي حجم العينة.

الاعتبارات الأخلاقية

تم الحصول على الموافقة الأخلاقية من لجنة أخلاقيات البحث العلمي بجامعة الملك فيصل KFU-REC-2025-MAY-ETHICS3349، وكانت مشاركة أفراد العينة في الدراسة بصورة طوعية، مع الالتزام بسرية البيانات واستخدامها لأغراض البحث العلمي فقط.

أداة الدراسة

بعد مراجعة الأدبيات التربوية، والدراسات السابقة، وعديد من المصادر ذات الصلة بموضوع الدراسة، وفي مقدمتها المعايير والمسارات المهنية للمعلمين في المملكة العربية السعودية (هيئة تقويم التعليم والتدريب، 2017)، وكذا كتب الرخصة المهنية للمعلمين المتعلقة بالمعايير التربوية (القسم التربوي والقسم الكمي)، ومنها: كتاب العواجي (2023)، وكذا بعض الكتب في مجال القياس والتقويم، ومنها: كتاب الزيد وعليان (2005)، وكتاب عودة (2010). وكذلك، بعد استقراء كلٍّ من دراسة أبو جراد (2011)، ودراسة عودة (1990)، ودراسة النوافلة وآخرين (2021)، ودراسة Agu et al. (2013)، ودراسة Ogunsanmi et al. (2025) – بعد كلٍّ ما سبق ذكره – قامت الباحثة بإعداد مقياس، تكون في صورته الأولية من جزأين، وذلك على النحو الآتي: الجزء الأول: تضمن البيانات العامة، واشتمل على: الخدمة في التدريس، نوع المواد الدراسية، التخصص، طريقة تحليل نتائج الاختبارات التحصيلية، والجزء الثاني: تضمن 20 عبارة، تمثل مهارات تحليل نتائج الاختبارات التحصيلية، الواجب توافرها لدى معلمات التعليم العام، ويتم الاستجابة عنها على مقياس ليكرت الخماسي بإحدى الاستجابات الخمس الآتية: مرتفعة جداً، مرتفعة، متوسطة، منخفضة، ومنخفضة جداً، وتأخذ هذه الاستجابات

الخمس الدرجات الخمس الآتية "5، 4، 3، 2، 1" على الترتيب، وجميع عبارات المقياس موجبة الاتجاه؛ أي: إنه كلما ارتفعت الدرجة الكلية على المقياس، أشار ذلك إلى ارتفاع درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية.

ويمتد مدى الدرجات على عبارات المقياس من 1 إلى 5؛ لذلك حوّلت درجة كل معلمة على المقياس إلى 5 درجات عن طريق القسمة على عدد عبارات المقياس وهو 18 عبارة، ثم صُنّفت تلك الإجابات إلى 5 درجات أو فئات متساوية في المدى؛ وذلك من خلال المعادلة الآتية:

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

$$= (5 - 1) \div 5 = 0.80, \text{ لنحصل على التصنيف الموضح في جدول 2.}$$

جدول 2

توزيع مدى الاستجابة وفق التدرج المستخدم في المقياس

الاستجابة	مدى الاستجابة	درجة الممارسة
مرتفعة جداً	من 4.20 إلى 5	مرتفعة جداً
مرتفعة	من 3.40 لأقل من 4.20	مرتفعة
متوسطة	من 2.60 لأقل من 3.40	متوسطة
منخفضة	من 1.80 لأقل من 2.60	منخفضة
منخفضة جداً	من 1 لأقل من 1.80	منخفضة جداً

ثبات الأداة

حُسب ثبات عبارات المقياس؛ بتطبيقها على عينة مكونة من 42 معلمة، تم اختيارهن عشوائياً من العينة الأساسية للدراسة، وذلك بطريقتين، هما: حساب معامل ألفا كرونباخ Alpha-Cronbach (بعدد عبارات المقياس): وفي كل مرة، كانت تحذف درجات إحدى العبارات من الدرجة الكلية للمقياس. وتحسب معاملات الارتباط بين درجات العبارة والدرجات الكلية للمقياس: باستخدام معامل ارتباط بيرسون (الاتساق الداخلي).

أما الثبات الكلي للمقياس؛ فحُسب بطريقتين، الأولى: عن طريق معامل ألفا كرونباخ. الثانية: عن طريق معامل الثبات؛ بطريقة التجزئة النصفية سبيرمان-براون (Spearman-Brown).

صدق الأداة

قيس صدق الأداة من خلال:

- صدق المحكمين: عرضت على مجموعة من المحكمين المتخصصين في المناهج وطرق التدريس، وعلم النفس التربوي، ومعلمات في التعليم العام؛ للحكم على مدى ملائمة العبارات لما وضعت لقياسه، وكفايتها، ومناسبتها، وصياغتها اللغوية، والتعديل المقترح. وفي ضوء ملاحظاتهم، دُمجت العبارة 11 مع العبارة 12 لتصبحا عبارة واحدة فقط. وكذلك، دُمجت العبارة 18 مع العبارة 19 لتصبحا عبارة واحدة فقط. ومن ثم؛ أصبح المقياس جاهزاً للتطبيق، ويتكون من 18 عبارة.
- صدق العبارات: حُسب صدق عبارات المقياس بتطبيقها على عينة استطلاعية مكونة من 42 معلمة، تم اختيارهن عشوائياً من العينة الأساسية للدراسة؛ وذلك عن طريق حساب معامل الارتباط بين درجة العبارة والدرجة الكلية للمقياس، عند حذف درجة العبارة من الدرجة الكلية للمقياس؛ باعتبار أن بقية عبارات المقياس محكٌ للعبارة (حسن، 2018؛ أبو هاشم، 2004). كما حُسب الصدق الكلي للمقياس؛ عن طريق حساب معامل الصدق الذاتي للمقياس، الذي يساوي الجذر التربيعي لمعامل ثبات المقياس. ويوضح جدول 3 معاملات ثبات وصدق عبارات المقياس على العينة الاستطلاعية.

جدول 3

معاملات ثبات عبارات المقياس وصدقها (ن = 42)

العبارة	معامل ألفا كرونباخ Alpha-Cronbach	معامل ارتباط درجة العبارة بالدرجة الكلية للمقياس (ثبات)	معامل ارتباط درجة العبارة بالدرجة الكلية للمقياس عند حذف درجة العبارة (صدق)
1	0.952	**0.55	**0.50
2	0.952	**0.52	**0.47
3	0.953	**0.44	**0.39
4	0.953	**0.45	**0.40
5	0.954	**0.41	**0.35
6	0.948	**0.78	**0.75

تابع/ جدول 3

معاملات ثبات عبارات المقياس وصدقها (ن = 42)

العبرة	معامل ألفا كرونباخ Alpha-Cronbach	معامل ارتباط درجة العبرة بالدرجة الكلية للمقياس (ثبات)	معامل ارتباط درجة العبرة بالدرجة الكلية للمقياس عند حذف درجة العبرة (صدق)
7	0.946	**0.87	**0.85
8	0.950	**0.70	**0.66
9	0.948	**0.79	**0.76
10	0.949	**0.72	**0.69
11	0.948	**0.78	**0.75
12	0.949	**0.74	**0.70
13	0.947	**0.85	**0.82
14	0.945	**0.92	**0.90
15	0.945	**0.91	**0.90
16	0.946	**0.88	**0.85
17	0.946	**0.88	**0.86
18	0.945	**0.94	**0.92

معامل Alpha-Cronbach الكلي للمقياس = 0.952

معامل الثبات بطريقة التجزئة النصفية ل Spearman-Brown = 0.980

معامل الصدق الكلي (الذاتي) للمقياس = 0.976

ملاحظة. ** دال إحصائياً عند مستوى (0.01).

يتضح من جدول 3:

- أن جميع معاملات ألفا كرونباخ Alpha-Cronbach عند حذف أي عبارة من عبارات المقياس، أقل من أو يساوي معامل ألفا كرونباخ العام للمقياس، في حالة وجود جميع عبارات المقياس؛ أي إن وجود أي عبارة من عبارات المقياس لا يؤدي إلى انخفاض معامل الثبات الكلي للمقياس، وإن استبعاد أي عبارة من عباراته -أو حذفها- يؤدي إلى انخفاض معامل الثبات الكلي للمقياس، وهذا يشير إلى أن وجود كل عبارة يسهم بدرجة معقولة في الثبات الكلي للمقياس (حسن، 2018).

- أن جميع معاملات ارتباط درجة العبارة بالدرجة الكلية للمقياس دالة إحصائياً عند مستوى ($\alpha \leq 0.01$)؛ مما يدل على الاتساق الداخلي، وثبات جميع عبارات المقياس.
 - أن معاملي الثبات الكلي للمقياس بطريقتي (ألفا كرونباخ Alpha-Cronbach، والتجزئة النصفية Spearman-Brown) مرتفعان، ويعد معامل الثبات مرتفعاً إذا بلغ 0.80 فأكثر (أبو هاشم، 2004)؛ مما يدل على الثبات الكلي للمقياس.
 - أن جميع معاملات ارتباط درجة العبارة بالدرجة الكلية للمقياس -في حالة حذف درجة العبارة من الدرجة الكلية للمقياس- دالة إحصائياً عند مستوى ($\alpha \leq 0.01$)؛ مما يدل على صدق جميع عبارات المقياس.
 - أن معامل الصدق الكلي للمقياس مرتفع؛ مما يدل على الصدق الكلي للمقياس.
- ومن خلال الإجراءات السابقة، تأكد للباحثة ثبات المقياس وصدقه؛ ومن ثم صلاحيته لقياس درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية، وتشير الدرجة المرتفعة في المقياس إلى ارتفاع درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، في حين تشير الدرجة المنخفضة في المقياس إلى انخفاض درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، وكانت أعلى درجة يمكن أن تحصل عليها المعلمة في جميع عبارات المقياس هي 90 درجة، في حين كانت الدرجة 18 هي أقل درجة يمكن أن تحصل عليها المعلمة.

أسلوب تحليل البيانات

تمت معالجة البيانات وتحليلها عن طريق برنامج الحزم الإحصائية للعلوم الاجتماعية (SPSS)؛ وذلك باستخدام الأساليب الإحصائية الآتية: معامل ألفا كرونباخ (Cronbach's Alpha)، للتحقق من ثبات المقياس، ومعامل الارتباط بيرسون (Pearson Correlation Coefficient)؛ للتحقق من صدق عبارات المقياس وثباته. ومعامل الثبات بطريقة التجزئة النصفية سبيرمان/ براون (Spearman-Brown). والتكرارات، والنسبة المئوية لعدد المعلمات ذوات درجات الممارسة المختلفة. وكذلك المتوسط الحسابي، والانحراف المعياري، لكل عبارة من عبارات المقياس. في حين كان تحليل التباين في اتجاه واحد (One-Way ANOVA)؛ لبحث الفروق بين متوسطات درجة الممارسة في ضوء المتغيرات الآتية: الخدمة في التدريس، ونوع المواد الدراسية، والتخصص، واختبار كولموجوروف-سميرنوف (Kolmogorov-Smirnov)

للتحقق من اعتدالية التوزيع، واختبار ليفين (Levene) للتحقق من تجانس التباين في درجة الممارسة. ومربع إيتا الجزئي (Eta-Square) لحساب حجم التأثير.

النتائج

نتائج الإجابة عن السؤال الأول

نص السؤال الأول: ما مهارات تحليل نتائج الاختبارات التحصيلية الواجب توافرها لدى معلمات التعليم العام؟ وللإجابة عن هذا السؤال، صمّم مقياس مكون من 18 عبارة، تمثل مهارات تحليل نتائج الاختبارات التحصيلية، الواجب توافرها لدى معلمات التعليم العام، وقد أوضحت تلك المهارات بالتفصيل في الإجابة عن السؤال الثاني.

نتائج الإجابة عن السؤال الثاني

نص السؤال الثاني: ما درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؟ وللإجابة عن هذا السؤال، استخدم المتوسط الحسابي والنسبة المئوية للمتوسط بالنسبة إلى الدرجة الكلية، والتكرارات والنسبة المئوية لعدد المعلمات ذوات درجات الممارسة المختلفة، والمتوسط الحسابي والانحراف المعياري لكل عبارة من عبارات المقياس. وكانت النتائج على النحو الوارد في جدول 4، 5.

جدول 4

متوسط الدرجات والتكرارات والنسبة المئوية لعدد المعلمات ذوات درجات الممارسة المختلفة على المقياس، والنسبة المئوية للمتوسط (ن=105)

درجات ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية							
مرتفعة جداً		مرتفعة		متوسطة		منخفضة	
من 4.20 إلى 5		من 3.40 لأقل		من 2.60 لأقل		من 1.80 لأقل	
من 4.20		من 4.20		من 3.40		من 2.60	
العدد	النسبة	العدد	النسبة	العدد	النسبة	العدد	النسبة
30	%28.6	29	%27.6	28	%26.7	16	%15.2
2	%1.9						

المتوسط الكلي للمقياس * = 3.57

النسبة المئوية للمتوسط الكلي للمقياس = 71.4%

ملاحظة. * المتوسط الكلي للمقياس من 5 درجات.

يتضح من جدول 4:

- أن متوسط درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية لدى عينة الدراسة، من معلمات المرحلة الثانوية في كل التخصصات؛ قد بلغ 3.57 من 5، ونسبة هذا المتوسط 71.4% من الدرجة العظمى للمقياس، وهذا المتوسط يقع في مدى الاستجابة المرتفع، الذي يمتد من 3.40 لأقل من 4.20؛ مما يشير إلى أن درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية -بوجه عام- كانت بدرجة مرتفعة.
- أن أعلى نسبة من معلمات التعليم العام بالمرحلة الثانوية وقع متوسط ممارستهن لمهارات تحليل نتائج الاختبارات التحصيلية في درجة الممارسة المرتفعة جداً؛ إذ بلغت نسبتهن 28.6% من إجمالي عينة الدراسة، تلتها نسبة ذوات درجة الممارسة المرتفعة وقدرها 27.6%، ثم نسبة ذوات درجة الممارسة المتوسطة وقدرها 26.7%. أما نسبتا ذوات درجة الممارسة المنخفضة والمنخفضة جداً؛ فقد جاءتا في المرتبتين الأخيرتين، وقدرهما 15.2%، 1.9% من إجمالي عينة الدراسة. وهذا يشير إلى تدني نسبة المعلمات اللاتي يمارسن مهارات تحليل نتائج الاختبارات التحصيلية، سواء بصورة منخفضة أو منخفضة جداً.

جدول 5

المتوسط الحسابي، والانحراف المعياري، وترتيب كل عبارة من عبارات المقياس

م	العبارة	م	ع	الترتيب	درجة الممارسة
5	أحسب المتوسط الحسابي للدرجات.	4.34	1.00	1	مرتفعة جداً
4	أحسب مدى الدرجات (أكبر درجة - أقل درجة).	4.27	1.07	2	مرتفعة جداً
3	أوضح توزيع الدرجات بيانياً؛ مثل: مدرج تكراري، مضلع تكراري،....	4.13	1.14	3	مرتفعة
10	أستخدم النسبة المئوية لتحديد مستويات الطالبات.	3.95	1.27	4	مرتفعة
8	أحدد المنوال وهو الدرجة الأكثر تكراراً في توزيع الدرجات.	3.84	1.32	5	مرتفعة
11	أحسب معاملات السهولة والصعوبة لفقرات الاختبار.	3.82	1.36	6	مرتفعة
1	أرتب أوراق الاختبار بعد تصحيحها ترتيباً تصاعدياً أو تنازلياً بحسب الدرجة المستحقة.	3.75	1.34	7	مرتفعة

تابع/ جدول 5

المتوسط الحسابي، والانحراف المعياري، وترتيب كل عبارة من عبارات المقياس

م	العبارة	م	ع	الترتيب	درجة الممارسة
2	أرتب الدرجات في فئات بصورة جدول تكراري.	3.68	1.36	8	مرتفعة
6	أحسب وسيط الدرجات (الدرجة التي تقسم الدرجات بالتساوي، من حيث العدد).	3.66	1.36	9	مرتفعة
9	أحسب الرتبة المئينية لدرجة الطالبة، وهي: النسبة المئوية لدرجة الطالبة بالنسبة لدرجات الطالبات ككل؛ بمعنى: كم طالبة هي أعلى منهن؟ أو: كم طالبة هي أقل منهن؟	3.64	1.35	10	مرتفعة
12	أحسب معامل التمييز لفقرات الاختبار.	3.58	1.43	11	مرتفعة
7	أحسب التباين أو الانحراف المعياري للدرجات.	3.43	1.43	12	مرتفعة
13	أحسب فاعلية الموهبات (البدائل) في أسئلة الاختيار من متعدد.	3.42	1.41	13	مرتفعة
18	أحسب معامل الارتباط بين الدرجات؛ للكشف عن ثبات رتبة الطالبة في الصف.	3.15	1.43	14	متوسطة
14	أحسب معامل الثبات بطريقة الصور المتكافئة.	3.04	1.48	15	متوسطة
15	أحسب معامل الثبات بطريقة التجزئة النصفية.	2.98	1.47	16	متوسطة
16	أحسب معامل الثبات بطريقة الاتساق الداخلي.	2.84	1.44	17	متوسطة
17	أحسب اعتدالية توزيع الدرجات؛ من خلال معاملات الالتواء والتفلطح.	2.74	1.51	18	متوسطة
	الدرجة الكلية لممارسة المعلمات لتحليل الاختبارات التحصيلية	3.57	0.97	-	مرتفعة

يتضح من جدول 5:

- أن المقياس تكوّن من 18 عبارة، منها (2) عبارتان بدرجة ممارسة مرتفعة جداً، و(11) إحدى عشرة عبارة جاءت بدرجة ممارسة مرتفعة، و(5) خمس عبارات جاءت بدرجة ممارسة متوسطة. ولا توجد أي عبارة بدرجة ممارسة منخفضة، أو منخفضة جداً، من وجهة نظر عينة الدراسة، المكونة من معلمات التعليم العام بالمرحلة الثانوية.

- أن أعلى عبارات المقياس، هي العبارة: "أحسب المتوسط الحسابي للدرجات"؛ وأن أدناها هي العبارة: "أحسب اعتدالية توزيع الدرجات من خلال معاملات الالتواء والتفلطح"، وأن الفرق بين أعلى متوسط وأقل متوسط هو "4.34 - 2.74 = 1.60"، وهو فرقٌ كبيرٌ نسبياً؛ مما يدل على وجود تفاوت -إلى حد كبير- في درجات ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية المتضمنة في جميع عبارات المقياس.
 - أن قيم الانحرافات المعيارية لعبارات المقياس امتدت من 1 إلى 1.51، وهي قيم تشير إلى وجود تشتت -أي تباين- في استجابات العينة، عن متوسط عبارات المقياس، وخاصة في حالة العبارات التي حظيت بالرتب الأخيرة، من وجهة نظر معلمات التعليم العام بالمرحلة الثانوية.
- وجاء ترتيب العبارات بصورة تنازلية، وفقاً للآتي:
- جاءت العبارة: "أحسب المتوسط الحسابي للدرجات" في المرتبة الأولى، بمتوسط حسابي 4.34، تلتها في المرتبة الثانية العبارة: "أحسب مدى الدرجات (أكبر درجة - أقل درجة)، بمتوسط حسابي قدره 4.27، ويقع هذان المتوسطان في مدى الاستجابة مرتفعة جداً؛ الذي يمتد من 4.20 إلى 5؛ مما يشير إلى أن معلمات التعليم العام بالمرحلة الثانوية يمارسن حساب المتوسط الحسابي للدرجات ومدى الدرجات، بدرجة مرتفعة جداً.
 - أن العبارات التي احتلت المرتبة من الثالثة إلى الثالثة عشرة، هي العبارات الإحدى عشرة الآتية: عبارة: "أوضح توزيع الدرجات بيانياً؛ مثل: مدرج تكراري، مضلع تكراري، ..."، وعبارة: "أستخدم النسبة المئوية لتحديد مستويات الطالبات"، وعبارة: "أحدد المنوال، وهو الدرجة الأكثر تكراراً في توزيع الدرجات"، وعبارة: "أحسب معاملات السهولة والصعوبة لفقرات الاختبار"، وعبارة: "أرتب أوراق الاختبار، بعد تصحيحها ترتيباً تصاعدياً أو تنازلياً، بحسب الدرجة المستحقة"، وعبارة: "أرتب الدرجات في فئات بصورة جدول تكراري"، وعبارة: "أحسب وسيط الدرجات (الدرجة التي تقسم الدرجات بالتساوي، من حيث العدد)"، وعبارة: "أحسب الرتبة المئينية لدرجة الطالبة، وهي: (النسبة المئوية لدرجة الطالبة بالنسبة لدرجات الطالبات ككل؛ بمعنى: كم طالبة هي أعلى منهن؟، أو: كم طالبة هي أقل منهن؟)"، وعبارة: "أحسب معامل التمييز لفقرات الاختبار"، وعبارة: "أحسب التباين أو الانحراف المعياري للدرجات"، وعبارة: "أحسب فاعلية المموهات (البدائل) في أسئلة الاختيار من متعدد"؛ وذلك بمتوسطات 4.13، 3.95، 3.84، 3.82، 3.75، 3.68، 3.66.

- 3.64، 3.58، 3.43، 3.42 على الترتيب، وجميع هذه المتوسطات تقع في مدى الاستجابة "مرتفعة"؛ مما يشير إلى أن معلمات التعليم العام بالمرحلة الثانوية يمارسن مهارات تحليل نتائج الاختبارات التحصيلية المتضمنة في هذه العبارات الإحدى عشرة بدرجة مرتفعة.
- أن العبارات التي احتلت المرتبة من الرابعة عشرة إلى الثامنة عشرة (الأخيرة) هي العبارات الآتية: عبارة: "أحسب معامل الارتباط بين الدرجات للكشف عن ثبات رتبة الطالبة في الصف"، وعبارة: "أحسب معامل الثبات بطريقة الصور المتكافئة"، وعبارة: "أحسب معامل الثبات بطريقة التجزئة النصفية"، وعبارة: "أحسب معامل الثبات بطريقة الاتساق الداخلي"، وعبارة: "أحسب اعتدالية توزيع الدرجات من خلال معاملات الالتواء والتفلطح"؛ وذلك بمتوسطات 3.15، 3.04، 2.98، 2.84، 2.74 على الترتيب، وجميع هذه المتوسطات تقع في مدى الاستجابة "متوسطة"؛ مما يشير إلى أن معلمات التعليم العام بالمرحلة الثانوية يمارسن مهارات تحليل نتائج الاختبارات التحصيلية المتضمنة في هذه العبارات الخمس بدرجة متوسطة.

مناقشة نتائج السؤال الثاني وتفسيراتها

أظهرت نتائج السؤال الثاني:

- أن درجة ممارسة معلمات التعليم العام بالمرحلة الثانوية لمهارات تحليل نتائج الاختبارات التحصيلية -بوجه عام- كانت بدرجة مرتفعة؛ إذ بلغ متوسط ممارستهن 3.57 من 5.
- أن أعلى نسبة من معلمات التعليم العام بالمرحلة الثانوية، وقع متوسط ممارستهن لمهارات تحليل نتائج الاختبارات التحصيلية في درجة الممارسة المرتفعة جداً؛ إذ بلغت نسبتهن 28.6%، تلتها نسبة ذوات درجة الممارسة المرتفعة، وقدرها 27.6%؛ ثم نسبة ذوات درجة الممارسة المتوسطة؛ وقدرها 26.7%. أما نسبتا ذوات درجة الممارسة المنخفضة والمنخفضة جداً، فقد جاءتا في المرتبتين الأخيرتين، وقدرهما 15.2% و 1.9%، من إجمالي عينة الدراسة.

ويمكن تفسير هذه النتيجة؛ بأن كثيراً من المعلمات يعتمدن في تحليل النتائج على مؤشرات التحليل الكلي للاختبار؛ مثل: حساب المتوسط الحسابي للدرجات، والانحراف المعياري، ونسب النجاح والرسوب، إضافة إلى تحديد المستوى العام لصعوبة الاختبار؛

فهذا النوع من التحليل يُعد من الممارسات الشائعة في البيئة المدرسية؛ نظراً لسهولة تطبيقه وسرعته، من خلال تحليل نتائج الدرجات عبر الأنظمة الإلكترونية المستخدمة في المدارس "برنامج نور"؛ إذ إن معظم معلمات عينة الدراسة كن يقمن بتحليل نتائج الاختبارات التحصيلية باستخدام برامج إلكترونية بنسبة 94.29%، من إجمالي حجم العينة. ولذلك قد تشعر المعلمات بأنهن يقمن بعملية تحليل لنتائج الاختبارات بصورة مستمرة، الأمر الذي ينعكس في ارتفاع تقديرهن لدرجة ممارستهن لهذه العملية.

كما أن هذا الارتفاع يرتبط في كثير من الأحيان بالتركيز على التحليل الكلي للاختبار، دون التعمق في تحليل خصائص فقرات الاختبار نفسها؛ مثل معامل الصعوبة ومعامل التمييز وفاعلية البدائل. ويترتب على ذلك أن بعض المعلمات قد يكتفين بتفسير النتائج في ضوء الدرجات الإجمالية أو المتوسطات العامة، دون إجراء تحليل تفصيلي يكشف عن جودة الفقرات ومدى قدرتها على قياس الأهداف التعليمية بدقة.

وبناءً على ذلك؛ يمكن القول إن ارتفاع درجة الممارسة قد يعكس ممارسة شكلية أو عامة لعملية تحليل النتائج، تقوم أساساً على المؤشرات الكلية للأداء، في حين أن التحليل الجزئي الخاص بالفقرات قد لا يُمارس بالدرجة نفسها. وقد يرجع ذلك إلى عدة عوامل، منها: محدودية التدريب المتخصص في مجال تحليل فقرات الاختبارات، أو ضيق الوقت المتاح للمعلمات، أو التركيز في الممارسات المدرسية على متابعة النتائج العامة أكثر من تحليل بنية الاختبار نفسها.

لذلك، اختلفت النتائج في هذه الدراسة مع نتائج الدراسات الآتية: دراسة أبو جراد (2011)، ودراسة النوافلة وآخرين (2021)، ودراسة Yusrizal et al. (2017)، التي أظهرت أن درجة ممارسة المعلمين لتحليل نتائج الاختبارات التحصيلية كانت بدرجة منخفضة. ودراسة عدائكة وآخرين (2018)، التي أوضحت أن مستوى معرفة أساتذة التعليم الثانوي لكفايات التحليل والتطبيق في إعداد الاختبارات التحصيلية، كانت منخفضة، ودراسة المسعودي والجاسر (2018)، التي أظهرت أن كفاية تفسير نتائج عملية القياس والتقويم، التي تحتوي على بعض الممارسات لتحليل الاختبار كانت بدرجة قليلة.

بينما اتفقت نتيجة هذه الدراسة، مع دراسة Kissi et al. (2023)، التي أظهرت وجود فجوة كبيرة بين الكفاءة الذاتية المدركة للمعلمين في بناء الاختبارات وجودة التطبيق الفعلي؛ ففي هذه الدراسة أشارت المعلمات إلى أنهن يمارسن مهارات تحليل نتائج الاختبارات التحصيلية بكفاءة، وأن متوسط درجة الممارسة قد بلغ 3.57 من 5، ونسبة هذا المتوسط 71.4% من الدرجة العظمى للمقياس، ويقع في مدى الاستجابة المرتفعة. إلا أن الدراسة كشفت وجود فجوة أو خلل بين الكفاءة الذاتية المدركة للمعلمات في ممارسة مهارات تحليل نتائج الاختبارات التحصيلية والتطبيق الفعلي لتحليل نتائج الاختبارات التحصيلية؛ إذ أظهر الاطلاع على عينة من تحليل نتائج الاختبارات التحصيلية التي قامت بها المعلمات في نهاية العام الدراسي عكس ذلك؛ فقد قامت الباحثة بفحص ما يقارب 60 تحليلًا لـ 60 معلمة في أغلب التخصصات؛ فكان التركيز على بعض المعالجات الإحصائية في التحليل الكلي -بشكل عام- دون أي إجراء للتحليل الجزئي، الذي يتناول تحليل فقرات الاختبار، وهذا ما سبق توضيحه.

وقد تُعَلَّل الباحثة ذلك، بأن المعلمات كن لا يقمن بهذه الممارسات بشكل صحيح وكما هو مطلوب؛ فمثلاً حساب معامل الصعوبة لفقرات الاختبار كان يتم بصورة تقديرية، وليس كما هو مطلوب، وهو ما أكدته جخراب وآخرون (2015) من أن مثل هذه الكفايات غير واضحة المعالم لدى المعلمين؛ فهي ممارسات عشوائية، وليس ككفاية فعلية ممارسة بشكل صحيح. وقد أشارت دراسة الزبون (2013) إلى أن المعلمين في حاجة إلى عقد دورات تدريبية، فيما يخص تحليل نتائج الاختبارات بالطرق الإحصائية، وأن أهم المعوقات التي تحول دون بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد، تتعلق بعدم القدرة على تحليل نتائج الاختبارات وتفسيرها.

كما أظهرت نتائج السؤال الثاني:

- أن أعلى عبارات المقياس؛ من حيث درجة الممارسة، هي العبارة: "أحسب المتوسط الحسابي للدرجات"، وهي في درجة ممارسة "مرتفعة جداً"، وأن أدناها هي العبارة: "أحسب اعتدالية توزيع الدرجات من خلال معاملات الالتواء والتفلطح"، وهي في درجة ممارسة "متوسطة". وأن الفرق بين أعلى متوسط وأقل متوسط هو $4.34 - 2.74 = 1.60$ ، وهو فرق كبير نسبياً؛ مما يدل على وجود تفاوت -إلى حد كبير- في درجات الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية؛ المتضمنة في جميع عبارات المقياس.

ويمكن تفسير نتيجة أعلى عبارات المقياس: "أحسب المتوسط الحسابي للدرجات"، في ضوء طبيعة المتوسط الحسابي بوصفه من أبسط وأشهر مقاييس النزعة المركزية المستخدمة في تحليل البيانات التربوية؛ إذ يوفر مؤشراً عاماً وسريعاً عن مستوى أداء الطلبة في الاختبار، بالإضافة إلى ما ذكر آنفاً عن توافر البرامج الإلكترونية التي ساعدت في تسهيل العمليات الحسابية، ؛ فهي بذلك تتفق مع دراسة النوافلة وآخرين (2021)، في أن حساب المتوسط الحسابي من الممارسات السهلة جداً، ويستطيع جميع المعلمين تنفيذها بسهولة، ولا تحتاج إلى إحصائيات متقدمة. في حين تعزو الباحثة أدنى عبارات المقياس: "أحسب اعتدالية توزيع الدرجات من خلال معاملات الالتواء والتفلطح"، إلى حاجة المعلمات للتدريب المناسب على هذه الممارسات؛ لتحسين مستواه وتطوير مهارتهن.

وعلى الرغم من وجود هذه الأساليب الإحصائية في القسم الكمي لاختبارات الرخصة المهنية، فإن الدورات التدريبية الخاصة بالرخص المهنية كانت تُدرب المعلمين على دخول الاختبارات فقط؛ فالمعلمون والمعلمات في حاجة إلى برامج تدريبية تركز على الجوانب التطبيقية في تحليل نتائج الاختبارات التحصيلية، وقد أوصت بذلك عديد من الدراسات، منها: دراسة أبو جراد (2011)، ودراسة الزبون (2013)، ودراسة المسعودي والجاسر (2018)، ودراسة Sengkaton et al. (2020).

وقد أظهرت نتائج دراسة أبو جراد (2011)، ودراسة النوافلة وآخرين (2021) أن حساب معاملي الالتواء والتفلطح كان يُمارس بصورة منخفضة من قبل المعلمين. كما يمكن تفسير الفرق -وهو فرق كبير نسبياً- في مستويات الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية بأنه نتيجة لاختلاف رتبهن؛ فقد أوضحت المعايير والمسارات المهنية للمعلمين في المملكة العربية السعودية وصفاً لأداء المعلمين، بحسب رتبهن في مجال التقويم.

نتائج الإجابة عن السؤال الثالث

نص السؤال الثالث: هل تختلف درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ باختلاف كل من: الخدمة في التدريس، نوع المواد الدراسية، التخصص؟ وللإجابة عن هذا السؤال، صيغت الفروض الآتية:

1 - توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$)؛ بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير الخدمة في التدريس.

2 - توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$)؛ بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير نوع المواد الدراسية.

3 - توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$)؛ بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير التخصص.

وللتحقق من صحة الفروض الثلاثة السابقة؛ استخدم تحليل التباين في اتجاه واحد (One- Way ANOVA)، وفيما يلي اختبارات التحقق من صحة شروط استخدام تحليل التباين في اتجاه واحد (اعتدالية التوزيع، تجانس التباين)، وكانت النتائج على النحو الوارد في جدول 6.

جدول 6

نتائج اختبار كولموجوروف-سميرنوف (Kolmogorov-Smirnov) للتحقق من اعتدالية التوزيع ونتائج اختبار ليفين (Levene) للتحقق من تجانس التباين في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية بحسب المتغيرات الثلاثة

المتغير	المجموعة الفرعية للمتغير	اختبار كولموجوروف-سميرنوف			اختبار ليفين		
		قيمة الاختبار	درجات الحرية	مستوى الدلالة	قيمة الاختبار	درجات الحرية	مستوى الدلالة
الخدمة في التدريس	أقل من خمس سنوات	0.121	36	*0.20	0.789	2	102
	من 5 سنوات إلى أقل من 10 سنوات	0.173	23	0.07 غير دالة			
	من 10 سنوات فأكثر	0.129	46	0.06 غير دالة			

تابع/ جدول 6

نتائج اختبار كولموجوروف-سمير نوف (Kolmogorov-Smirnov) للتحقق من اعتدالية التوزيع ونتائج اختبار ليفين (Levene) للتحقق من تجانس التباين في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية بحسب المتغيرات الثلاثة

المتغير	المجموعة الفرعية للمتغير	اختبار كولموجوروف-سمير نوف			اختبار ليفين		
		قيمة الاختبار	درجات الحرية	مستوى الدلالة	قيمة الاختبار	درجات الحرية	مستوى الدلالة
نوع المواد الدراسية	علمية	0.115	55	0.06 غير دالة	0.427	2	102
	أدبية	0.104	39	0.20 غير دالة			
	متنوعة	0.123	10	0.15 غير دالة			
التخصص	اللغات	0.155	27	0.09 غير دالة	1.719	3	101
	العلوم الطبيعية	0.106	33	0.20 غير دالة			
	الرياضيات والمهارات الرقمية	0.129	23	0.12 غير دالة			
	الدراسات الإسلامية والعلوم الاجتماعية	0.114	22	0.20 غير دالة			

يتضح من جدول 6:

- أن قيم اختبار كولموجوروف-سمير نوف لجميع المجموعات الفرعية للمتغيرات الثلاثة (الخدمة في التدريس، نوع المواد الدراسية، التخصص) غير دالة إحصائياً؛ مما يشير إلى أن درجات ممارسة المعلمات لمهارات تحليل نتائج الاختبارات التحصيلية ذات توزيع اعتدالي لدى المجموعات الفرعية لهذه المتغيرات الثلاثة.
- أن قيم اختبار ليفين للمتغيرات الثلاثة (الخدمة في التدريس، نوع المواد الدراسية، التخصص) غير دالة إحصائياً؛ مما يشير إلى أن درجات ممارسة المعلمات لمهارات

تحليل نتائج الاختبارات التحصيلية ذات تباين متجانس في المجموعات الفرعية للمتغيرات المستقلة الثلاثة.

مما سبق تم التحقق من صحة شروط استخدام تحليل التباين في اتجاه واحد، وفيما يلي نتائج الفروض الثلاثة ومناقشتها.

نتائج الفرض الأول ومناقشتها

للتحقق من صحة الفرض الأول، استخدم تحليل التباين في اتجاه واحد One- Way ANOVA؛ لبحث الفروق بين متوسطات درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير الخدمة في التدريس، ومن ناحية أخرى حُسبت متوسطات درجة الممارسة للمتغير نفسه، ومربع إيتا (Eta-Square) لحساب حجم التأثير، وكانت النتائج على النحو الوارد في جدول 7.

جدول 7

نتائج تحليل التباين في اتجاه واحد، عند دراسة الفروق في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، التي تُعزى لمتغير الخدمة في التدريس، ومتوسطات درجة الممارسة للمتغير، ومربع إيتا

سنوات الخدمة في التدريس	ن	م	ع	قيمة (ف)	مستوى الدلالة	مربع إيتا
أقل من خمس سنوات	36	3.50	0.91	1.06	0.35	0.02
من 5 سنوات إلى أقل من 10 سنوات	23	3.39	0.88		غير دالة	
من 10 سنوات فأكثر	46	3.72	1.05			

يتضح من جدول 7:

- عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$)؛ بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير الخدمة في التدريس؛ إذ إن قيمة (ف) تساوي 1.06، وهي قيمة غير دالة إحصائياً؛ مما يدل على عدم تحقق صحة الفرض الأول.
- وجود تقارب بين متوسطات درجة ممارسة معلمات التعليم العام ذوات سنوات الخدمة المختلفة في التدريس في مهارات تحليل نتائج الاختبارات التحصيلية.

- أن قيمة مربع إيتا تساوي 0.02؛ وتعني أن متغير الخدمة في التدريس يفسر 2% من التباين في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، وهي نسبة تشير إلى حجم تأثير ضعيف لمتغير الخدمة في التدريس. وهذا يدل على وجود اتفاق بين الدلالة الإحصائية والدلالة العملية المتمثلة في حجم التأثير فيما يتعلق بتأثير متغير الخدمة في التدريس.

ويمكن تفسير هذه النتائج بأن المعلمات يخضعن للظروف التدريبية نفسها، خاصة بعد إغلاق مكاتب الإشراف التربوي. كما ترى الباحثة، في ضوء اعتماد المعلمات على التحليل الكلي لنتائج الاختبارات، أن ممارسة مهارات تحليل النتائج أقل تأثراً بخبرة المعلمة أو طول سنوات الخدمة؛ لأنها تعتمد على بيانات عامة وسهلة الاستخدام يمكن لأي معلمة التعامل معها، سواء كانت حديثة الخدمة أم ذات خبرة طويلة. كما تعزو الباحثة هذه النتيجة إلى أن هذا النمط من الممارسة يعكس تركيز المدارس على مؤشرات عامة لمتابعة تحصيل الطالبات؛ مما يجعل العملية موحدة نسبياً لجميع المعلمات، ويقلل من أثر طول الخدمة على مستوى الممارسة.

وهي بذلك تتفق مع دراسة أبو جراد (2011)، ودراسة النوافلة وآخرين (2021)؛ من حيث عدم وجود أثر للخبرة التدريسية على ممارسة المعلمين لتحليل نتائج الاختبارات التحصيلية، ودراسة الزبون (2013)؛ التي أوضحت أن عامل خبرة المعلم لا يؤثر في مجال إتقانه للمهارات المتعلقة ببناء الاختبارات التحصيلية، وتختلف مع دراسة المسعودي والجاسر (2018)، في أنه توجد فروق في ممارسة المعلمات لكفايات القياس والتقويم تعود إلى الخبرة.

نتائج الفرض الثاني ومناقشتها

للتحقق من صحة الفرض الثاني، استخدم تحليل التباين في اتجاه واحد One- Way ANOVA؛ لبحث الفروق بين متوسطات درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير نوع المواد الدراسية، ومن ناحية أخرى حُسبت متوسطات درجة الممارسة للمتغير نفسه، ومربع إيتا (Eta-Square) لحساب حجم التأثير، وكانت النتائج على النحو الوارد في جدول 8.

جدول 8

نتائج تحليل التباين في اتجاه واحد، عند دراسة الفروق في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، التي تُعزى لمتغير نوع المواد الدراسية، ومتوسطات درجة الممارسة للمتغير، ومربع إيتا

نوع المواد الدراسية	ن	م	ع	قيمة (ف)	مستوى الدلالة	مربع إيتا
علمية	56	3.59	0.93	0.07	0.94	0.001
أدبية	39	3.56	1.02		غير دالة	
(علمية وأدبية)	10	3.47	1.06			

يتضح من جدول 8:

- عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$)؛ بين متوسطات درجة ممارسة المعلمات لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير نوع المواد الدراسية؛ إذ إن قيمة (ف) تساوي 0.07، وهي قيمة غير دالة إحصائياً؛ مما يدل على عدم تحقق صحة الفرض الثاني.
- وجود تقارب بين متوسطات درجة ممارسة المعلمات؛ اللاتي كن يُدرسن مواد دراسية مختلفة، في مهارات تحليل نتائج الاختبارات التحصيلية.
- أن قيمة مربع إيتا تساوي 0.001؛ وتعني أن متغير نوع المواد الدراسية يفسر 0.1% من التباين في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، وهي نسبة تشير إلى حجم تأثير ضعيف جداً لمتغير نوع المواد الدراسية. وهذا يدل على وجود اتفاق بين الدلالة الإحصائية والدلالة العملية المتمثلة في حجم التأثير فيما يتعلق بتأثير متغير نوع المواد الدراسية.

ويمكن تفسير هذه النتيجة بما ذكر في مناقشة الفرض الأول، بالإضافة إلى اعتماد المدارس على إجراءات ونماذج موحدة لتحليل نتائج الاختبارات؛ إذ تُطبّق هذه النماذج على جميع المعلمات دون تمييز بين التخصصات المختلفة؛ مما يؤدي إلى ممارسة المعلمات لعملية التحليل بأسلوب متقارب، وهذا يفسر عدم ظهور فروق في درجة الممارسة تعزى لاختلاف المادة الدراسية. إضافة إلى مجتمعات التعلم المهنية ودورها في تنمية الكفايات التدريسية لدى

المعلمات، وهو ما أكدته دراسة الحارثي والغامدي (2025)، ودراسة العتيبي (2022). ونظراً لأن هذه الدراسة تختلف عن الدراسات السابقة في بعض المتغيرات، فإن هذا المتغير غير موجود في الدراسات السابقة.

نتائج الفرض الثالث ومناقشتها

للتحقق من صحة الفرض الثالث؛ استخدم تحليل التباين في اتجاه واحد One- Way ANOVA؛ لبحث الفروق بين متوسطات درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير التخصص، ومن ناحية أخرى حُسبت متوسطات درجة الممارسة للمتغير نفسه، ومربع إيتا لحساب حجم التأثير، وكانت النتائج على النحو الوارد في جدول 9.

جدول 9

نتائج تحليل التباين في اتجاه واحد، عند دراسة الفروق في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، التي تُعزى لمتغير التخصص، ومتوسطات درجة الممارسة للمتغير، ومربع إيتا

النخصص	ن	م	ع	قيمة (ف)	مستوى الدلالة	مربع إيتا
اللغات	27	3.47	1.00	1.05	0.37	0.03
العلوم الطبيعية	33	3.45	0.97		غير دالة	
الرياضيات والمهارات الرقمية	23	3.88	0.80			
الدراسات الإسلامية والعلوم الاجتماعية	22	3.56	1.08			

يتضح من جدول 9:

- عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$)؛ بين متوسطات درجة ممارسة معلمات التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية؛ تُعزى لمتغير التخصص؛ إذ إن قيمة (ف) تساوي 1.05، وهي قيمة غير دالة إحصائياً؛ مما يدل على عدم تحقق صحة الفرض الثالث.
- وجود تقارب بين متوسطات درجة ممارسة المعلمات ذوات التخصصات المختلفة في مهارات تحليل نتائج الاختبارات التحصيلية.

- أن قيمة مربع إيتا تساوي 0.03، وتعني أن متغير التخصص يفسر 3% من التباين في درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، وهي نسبة تشير إلى حجم تأثير ضعيف لمتغير التخصص. وهذا يدل على وجود اتفاق بين الدلالة الإحصائية والدلالة العملية المتمثلة في حجم التأثير فيما يتعلق بتأثير متغير التخصص.

ويمكن تفسير هذه النتيجة، بما ذكر في مناقشة الفرضين الأول والثاني. ونظراً لأن هذه الدراسة تختلف عن الدراسات السابقة، في تناولها درجة الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية بصورة مباشرة ومفصلة للتحليلين الكلي والجزئي؛ وكذلك في بعض متغيرات الدراسة؛ فهي بذلك تتفق مع دراسة المسعودي والجاسر (2018)، في عدم وجود فروق في ممارسة المعلمات لكفايات القياس والتقويم، تعود إلى التخصص.

الخلاصة

استنتاجات الدراسة

يمكن عرض أهم الاستنتاجات، وذلك على النحو الآتي:

- أن درجة ممارسة المعلمات بالمرحلة الثانوية لمهارات تحليل نتائج الاختبارات التحصيلية - بوجه عام - كانت بدرجة مرتفعة؛ إذ بلغ متوسط ممارستهن 3.57 من 5.

- أن أعلى نسبة من معلمات التعليم العام بالمرحلة الثانوية، وقع متوسط ممارستهن لمهارات تحليل نتائج الاختبارات التحصيلية في درجة الممارسة المرتفعة جداً؛ إذ بلغت نسبتهن 28.6%، تلتها نسبة ذوات درجة الممارسة المرتفعة، وقدرها 27.6%؛ ثم نسبة ذوات درجة الممارسة المتوسطة؛ وقدرها 26.7%. أما نسبتا ذوات درجة الممارسة المنخفضة والمنخفضة جداً، فقد جاءتا في المرتبتين الأخيرتين، وقدرهما 15.2% و 1.9%، من إجمالي عينة الدراسة.

- أن أعلى عبارات المقياس، هي العبارة: "أحسب المتوسط الحسابي للدرجات"، وهي في درجة ممارسة مرتفعة جداً، وأنها هي العبارة: "أحسب اعتدالية توزيع الدرجات من خلال معاملات الالتواء والتفلطح"، وهي في درجة ممارسة متوسطة. وأن الفرق بين أعلى متوسط وأقل متوسط هو $4.34 - 2.74 = 1.60$ ، وهو فرق كبير نسبياً؛ مما يدل على وجود تفاوت -

إلى حد كبير - في درجات الممارسة لمهارات تحليل نتائج الاختبارات التحصيلية، المتضمنة في جميع عبارات المقياس.

- عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha \leq 0.05$): بين متوسطات درجة ممارسة المعلمات لمهارات تحليل نتائج الاختبارات التحصيلية؛ باختلاف كل من: الخدمة في التدريس، ونوع المواد الدراسية، والتخصص.

توصيات الدراسة

في ضوء نتائج الدراسة، توصي الباحثة بما يأتي:

- ضرورة تضمين برامج إعداد المعلمين في كليات التربية مقرر الإحصاء؛ بحيث يكون متزامناً مع مقرر التقويم التربوي.
- عقد دورات تدريبية تطبيقية للمعلمات -على اختلاف تخصصاتهن- في مجال التحليل الإحصائي لنتائج الاختبارات التحصيلية؛ لتحسين مهاراتهم في تحليل الاختبارات وتطويرها.
- تدريب المعلمات على إجراءات تقويمية؛ لمعرفة درجة مراعاة اختباراتهن وتحليلها لمعايير الاختبار الجيد.
- عقد ورش عمل لدراسة الصعوبات والتحديات التي تواجه المعلمات عند تحليل نتائج الاختبارات التحصيلية، ووضع الحلول لها؛ لتجويد تلك الاختبارات.
- ضرورة أن توفر المدرسة دعماً كافياً للمعلمين والمعلمات في مجال التطوير المهني، فيما يتعلق بالتقويم؛ من خلال تكوين وحدة للتحليل الإحصائي، تكون وظيفتها تحليل نتائج الاختبارات التحصيلية التي يعدها المعلمون، وتزويدهم بالتغذية الراجعة.
- عقد ورش عمل؛ لتعريف المعلمين والمعلمات بالبرامج الإلكترونية المستخدمة في إعداد جدول المواصفات للاختبار، وتحليل نتائجه، والمفاضلة بينها.

مقترحات الدراسة

تقترح الباحثة البحوث المستقبلية الآتية:

- إجراء بحث مماثل؛ لمعرفة درجة ممارسة معلمي التعليم العام لمهارات تحليل نتائج الاختبارات التحصيلية.

- إجراء بحث مقارنة بين درجة ممارسة المعلمين والمعلمات في التعليم العام؛ لمهارات تحليل نتائج الاختبارات التحصيلية.
- إجراء بحث؛ لمعرفة اتجاهات المعلمين والمعلمات نحو تحليل الاختبارات التحصيلية.
- إجراء بحث؛ لتحليل جودة أسئلة الاختبارات (التحليل التفصيلي) التي تعدها المعلمات في أي مادة من المواد الدراسية.
- تحديد الاحتياجات التدريبية في تحليل نتائج الاختبارات التحصيلية لمعلمات التعليم العام.

المراجع

- أبو جراد، حمدي يونس. (2011). مدى التزام المعلمين بتحليل نتائج الاختبارات التحصيلية وعلاقته باتجاهاتهم نحوها. *مجلة جامعة الأزهر بغزة*، 3(2)، 89-106.
- أبو فودة، باسل خميس، وبني يونس، نجاتي أحمد. (2012). *الاختبارات التحصيلية: مفهومها- كيفية إعدادها- أسس بنائها وتكوينها- وتطبيقات ميدانية*. دار المسيرة.
- أبو لبدة. سبيع محمد. (2008). *مبادئ القياس النفسي والتقييم التربوي*. دار الفكر.
- أبو هاشم، السيد. (2004). *الدليل الإحصائي في تحليل البيانات باستخدام SPSS*. مكتبة الرشد.
- الجامعة العربية المفتوحة. (2005). *القياس والتقويم وبناء الاختبارات المدرسية*. الجامعة العربية المفتوحة.
- جخراب، محمد عرفات، وجعفرور، ربيعة، وعدائكة، دنيا. (2015). كفايات إعداد الاختبارات التحصيلية الجيدة لدى أساتذة التعليم الثانوي. *مجلة العلوم الإنسانية*، 44(أ)، 515-540.
- الحارثي، هاجد دخيل محمد، والغامدي، علي عوض محمد. (2025). فاعلية مجتمعات التعلم المهنية في تنمية الكفايات التدريسية لدى معلمي اللغة العربية في المرحلة الابتدائية. *مجلة الآداب للدراسات النفسية والتربوية*، 7(2)، 400-452.
- حسن، عزت عبد الحميد محمد. (2018). *الإحصاء النفسي والتربوي تطبيقات باستخدام برنامج SPSS 18*. دار الفكر العربي.
- الزبون، حابس سعد. (2013). مدى كفاءة معلمي المدارس الحكومية الثانوية التابعة لإدارة التربية والتعليم بالطائف في بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد. *مجلة كلية التربية*، 14(1)، 92-130.
- الزبد، نادر فهد، وعليان، هشام عامر. (2005). *مبادئ القياس والتقويم في التربية* (ط.3). دار الفكر.
- صبري، ماهر إسماعيل. (2002). *الموسوعة العربية لمصطلحات التربية وتكنولوجيا التعليم*. مكتبة الرشد.

- عباس، محمد خليل، ونوفل، محمد بكر، والعبسي، محمد مصطفى، وأبو عواد، فريال محمد. (2009). *مدخل إلى مناهج البحث في التربية وعلم النفس* (ط.2). دار المسيرة.
- العتيبي، بدور ضيف الله. (2022). دور مجتمعات التعلم في تنمية الكفاءة المهنية لدى معلمات الكيمياء بمدينة الرياض. *المجلة العلمية لكلية التربية بجامعة أسيوط*, 38(10), 253-272.
- عدانكة، دنيا، وجخرا، محمد عرفات، وجعفرور، ربيعة. (2018). مستوى معرفة أساتذة التعليم الثانوي لكفايات إعداد الاختبارات التحصيلية الجيدة في ضوء أدوارهم الجديدة (دراسة استكشافية بثانويات ولاية الوادي وسط). *مجلة الباحث في العلوم الإنسانية والاجتماعية*, 33(3), 1055-1072.
- العدل. عادل محمد. (2017). *الإحصاء النفسي والتربوي تطبيقات على SPSS*. دار الكتاب الحديث.
- العساف، صالح حمد. (2006). *المدخل إلى البحث في العلوم السلوكية* (ط.4). مكتبة العبيكان.
- عطوي. جودت عزت. (2007). *أساليب البحث العلمي: مفاهيمه-أدواته-طرقه الإحصائية*. دار الثقافة للنشر والتوزيع.
- عطية. محسن علي. (2010). *البحث العلمي في التربية*. دار المناهج للنشر والتوزيع.
- علام. صلاح الدين محمود. (2006). *القياس والتقويم التربوي والنفسى أساسياته وتطبيقاته وتوجهاته المعاصرة*. دار الفكر العربي.
- العمر، عبد العزيز سعود. (2007). *لغة التربويين*. مكتب التربية العربي لدول الخليج.
- العواجي، محمد إبراهيم. (2023). *الرخصة المهنية التعليمية- المعايير التربوية* (ط.3). منصة ميم التعليمية.
- عودة، أحمد سليمان. (1990). تحليل نتائج الاختبارات المدرسية التي يعدها المعلمون بين الواقع والتطلعات. *مؤتة للبحوث والدراسات*, 5(1), 247-269.
- عودة، أحمد سليمان. (2010). *القياس والتقويم في العملية التدريسية*. دار الأمل للنشر والتوزيع.
- 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. *مجلة كلية التربية*, 34(164), 567-600.
- لفيل، كامل، وحمدان، فتحي. (2009). *الإحصاء STATISTICS*. دار المناهج للنشر والتوزيع.
- مراد. صلاح أحمد. (2000). *الأساليب الإحصائية في العلوم النفسية والتربوية والاجتماعية*. دار الأنجلو المصرية.
- المسعودي، أحمد، والجاسر، هند. (2018). كفايات القياس والتقويم لدى معلمة التعليم العام بتبوك وفق المعايير العالمية في ضوء بعض المتغيرات. *المجلة الدولية للدراسات التربوية والنفسية*, 4(2), 196-225.

- منسي. محمود عبد الحليم. (2007). *التقويم التربوي. دار المعرفة الجامعية*.
- النوافلة، وليد، والعمرى، علي، والمهيدات، رزان. (2021). درجة ممارسة معلمي العلوم لتحليل نتائج الاختبارات التحصيلية وعلاقته باتجاهاتهم نحوها. *دراسات للعلوم التربوية، 48* (1)، 541-523.
- هيئة تقويم التعليم والتدريب. (2017). *المعايير والمسارات المهنية للمعلمين في المملكة العربية السعودية*. <https://2u.pw/ImiLTVL>
- Agu, N. N., Onyekuba, C., & Anyichie, C. A. (2013). Measuring teachers' competencies constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill inventory. *Educ. Res. Rev, 8*(8), 431 -439.
- Ahmad, S. (2019). Analysis of test items used in an achievement test in physics secondary level. *Journal of Education and Practice, 10*(10), 90-96. <https://iiste.org/journals/index.php/JEP/article/view/47466>
- Aisyah, N. (2019, July 30-31). *An Analysis of the summative test on the 8th grade of Muhammadiyah Junior High School of Yogyakarta*. [Paper Presentation]. the Third International Conference on Sustainable Innovation –Humanity, Education and Social Sciences (IcoSI-HESS 2019). Yogyakarta, Indonesia. <https://doi.org/10.2991/icosihess-19.2019.86>
- Fatimah, S., Elzamzami, A. B., & Slamet, J. (2020). Item analysis of final test for the 9th grade students of SMPN 44 Surabaya in the academic year of 2019/2020. *JournEEL, 2*(1), 34-46.
- Kissi, P., Baidoo-Anu, D., Anane, E., & Annan-Brew, R. K. (2023). Teachers' test construction competencies in examination-oriented educational system: Exploring teachers' multiple-choice test construction competence. *Front. Educ, 8*:1154592. <https://doi.org/10.3389/feduc.2023.1154592>
- Manfaat, B., Nurazizah, A., & Misri, M. A. (2021). Analysis of mathematics test items quality for high school. *Jurnal Penelitian dan Evaluasi Pendidikan, 25*(1), 108-117.
- Ogunsanmi, O. A., Lijofi, V. A., & Gbolagun, A. E. (2025). Achievement testing competence of secondary school teachers in Osun State, Nigeria. *PAEDAGOGIA, 28*(1), 24-33.
- Professional Development Company. (2024). *Mastery-Grade* [Computer Software]. PDC. <https://mastery-grade.com>
- Sengkaton, I., Hambali, M., & Mirizon, S. (2020). Teacher-made summative test: An analysis of test format, Index of difficulty, discrimination, and distractors. *LINGUA: Jurnal Bahasa and Sastra, 20*(2), 152-158.

- Thompson, N.A., & Guyer, R. (2010). *User's manual for Iteman 4.1*. St. Paul MN: Assessment Systems Corporation.
- Tsegaye, K. N. (2019). Post exam item analysis: Implication for intervention. *bioRxiv*. <https://doi.org/10.1101/510081>
- Yusrizal, Y., Suliyana, S., & Basri, T. H. (2017). Analysis of knowledge, understanding and skills of physics teachers of state senior high schools in developing and analyzing test items. *Jurnal Pendidikan IPA Indonesia*, 6(2), 335-340.
- Zakariya, J. (2020). *Teachers' testing practices of achievement test in junior high schools in the Sissala East Municipality, Ghana* [Master's thesis, University of Cape Coast]. UCC Institutional Repository. <https://ir.ucc.edu.gh/xmlui/handle/123456789/4564>

د. وفاء شبيب محمد الدوسري، أستاذ المناهج وطرق التدريس المساعد بقسم المناهج وطرق التدريس، كلية التربية، جامعة الملك فيصل. دكتوراه في المناهج وطرق التدريس من جامعة الإمام محمد بن سعود الإسلامية بالمملكة العربية السعودية، 2017. عملت مشرفة تربوية للرياضيات في التعليم العام، ونالت جائزة التميز "قيادة متميزة" فئة المشرفات التربويات. الاهتمامات البحثية: المناهج وطرق التدريس، التربية العملية، التقويم التربوي، أبحاث الدماغ وتطبيقاتها في التعليم، والتكامل بين نظريات التعلم، إضافة إلى القضايا التعليمية المعاصرة.

waldossary@kfu.edu.sa