Dirasat: *Educational Sciences*, Volume 51, No. 2, 2024

Deanship of Scientific Research

The Level of Teachers Implementation of Quality Assessment Practices Standards in Public Schools in the Kingdom of Bahrain

Ahmed M. Alawadhi [1]* ***, Hassan Alomari*** [2]

[1]Ministry of Education, Manama, Bahrain

[2] Department of Educational Psychology, School of Educational Sciences, The University of Jordon, Amman, Jordon

Received: 31/10/2023
Revised: 11/12/2023
Accepted: 29/2/2024
Published: 15/6/2024

* *Corresponding author:*
AhmedAwadi77@gmail.com

Citation: Alawadhi , A. M. ., & Alomari , H. . (2024). The Level of Teachers Implementation of Quality Assessment Practices Standards in Public Schools in the Kingdom of Bahrain . *Dirasat: Educational Sciences*, *51*(2), 17–34.
https://doi.org/10.35516/edu.v51i2.6070

© 2024 DSR Publishers/ The University of Jordan.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC) license https://creativecommons.org/licenses/by-nc/4.0/

Abstract

Objectives: The current study aimed to investigate l the level of teachers' implementation of quality assessment practices in the Kingdom of Bahrain.

Methods: This study followed a descriptive analytical approach, and the study sample consisted of 430 teachers from various stages of education. The study used a stratified random sampling. To achieve the study objectives, an inventory of assessment practices quality standards was developed, consisting of 53 items distributed across five domains: development of assessment tools and methods, grading and score monitoring, reporting assessment results, ethical assessment practices, and utilization of assessment results. The psychometric properties of the final version of the list were verified.

Results: The study results revealed that teachers' implmemtation levels of quality assessment practices standards were high in all survey domains and overall. The means for the domains ranged from (3.08) to (3.44), with the domain of development of assessment tools and methods having the lowest mean, while the domain of ethical assessment practices had the highest mean. The study also indicated statistically significant differences at a significance level of (α=0.05) in the practice level attributed to the variables of qualification, favoring bachelor's degree holders, and undergoing training in measurement and evaluation, as well as those who received more specialized courses in test construction and measurement. However, no statistically significant differences at a significance level of (α=0.05) were found for the variables of gender, teaching stage, and years of experience.

Conclusions: The study revealed that teachers' practice levels of quality assessment practices standards were high. The study recommended the need for a similar study in private schools and the continuation of supporting and enhancing training policies and programs at the Ministry of Education to assess student achievement.

Keywords: Assessment practices, Practices of teacher's, Quality standards.

مستوى ممارسة المعلمين لمعايير جودة الممارسات التقييمية في المدارس الحكومية في مملكة البحرين

أحمد محمد العواضي [1]* ***، حسان العمري*** [2]

[1] وزارة التربية والتعليم، المنامة، مملكة البحرين

[2] قسم علم النفس التربوي، كلية العلوم التربوية، الجامعة الأردنية، عمّان، الأردن

ملخّص

الأهداف: هدفت الدراسة الحالية إلى الكشف عن مستوى ممارسة المعلّمين لمعايير جودة الممارسات التقييمية في مملكة البحرين.

المنهجية: اتبع في هذه الدراسة المنهج الوصفي التحليلي، وقد تكونت عينة الدراسة من (430) معلّما ومعلّمة من مختلف مراحل التعليم تم اختيارهم بالطريقة الطبقية العشوائية، ولتحقيق أهداف الدراسة؛ تم بناء قائمة معايير جودة الممارسات التقييمية مكونة من (53) فقرة في صورتها النهائيّة، موزعة على (5) مجالات، وهي (تطوير الأدوات والأساليب التقييمية، التصحيح ورصد الدرجات، الإخبار بنتائج التقويم، الممارسات الأخلاقية للتقويم، استخدام نتائج التقويم) بعد التحقق من خصائصها السايكومتريّة.

النتائج: وتبيّن من نتائج الدراسة أن مستويات ممارسة المعلّمين لمعايير جودة الممارسات التقييمية جاءت مرتفعة في جميع مجالات الاستبانة وعلى الاستبانة ككل، حيث تراوحت المتوسطات للمجالات بين (3.08 - 3.44)، حيث كان المجال الأقل متوسطًا هو مجال ممارسات تطوير الأدوات والأساليب التقييمية، بينما كان المجال الأعلى هو مجال الممارسات الأخلاقية للتقويم، واتضح من نتائج الدراسة وجودة فروقٍ ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) في مستوى الممارسة تعزى لمتغيري المؤهل ولصالح حاملي البكالوريوس، والخضوع لتدريب في القياس والتقويم ولصالح من تلقى أكثر من دورة متخصصة في بناء الاختبارات والقياس والتقويم، بينما لم تظهر أي فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) تعزى لمتغيرات الجنس والمرحلة التدريسية وسنوات الخبرة.

الخلاصة: اتضح أن مستويات ممارسة المعلمين لمعايير جودة الممارسات التقييمية جاء بدرجة مرتفعة، وأوصت الدراسة بضرورة بإجراء دراسة مماثلة على المدارس الخاصة، والاستمرار في دعم وتعزيز سياسات وبرامج التدريب بوزارة التربية والتعليم في تقييم تحصيل الطلبة.

الكلمات الدالة : الممارسات التقييمية، ممارسة المعلّمين، معايير الجودة.

المقدمة

تعد مهنة التعليم من المهن ذات المهام المتعددة، حيث يتعيّن على المعلم القيام بواجباته الوظيفية في مجالات العمل المدرسي المتنوعة، من حل المشاكل السلوكية، وإدارة سلوك الطلبة، وتنظيم البيئة الصفية، وتحضير وتقديم الدروس إضافة إلى عملية تقييم الطلبة التي تعدّ إحدى أكثر هذه المهام مركزيّة، وأهمية لما يترتب عليها من قرارات مصيرية بحق الطلبة.

ونظرًا لأهمية هذا التقييم وكون المعلّم مسؤولًا عن ذلك، قامت العديد من الجهود الدولية لإيجاد معايير تتعلق بممارسات المعلمين في تقييم تحصيل الطلبة عنونت بكفاءة المعلم في التقييم التعليمي للطلاب Teacher Competence، حيث تتطلب هذه المعايير مستوى مرتفعًا من المعارف والمهارات التي ينبغي اكتسابها عند تكوين المعلّم وتدريبه، حيث قامت مؤخرًا هذه الجهود بين اتحاد منظمات المعلمين الأمريكيين American Federations of Teachers (AFT)، والمجلس القومي للقياس التربوي National Council on Measurement of Education (NCME)، والجمعية القوميّة التربويّة National Education Association (NEA)، ثم استمرت الجهود حتى العام 2014 حين أصدرت الجمعية الأمريكية للبحوث التربوية The American Educational Research Association (AERA)، والمجلس القومي للقياس التربوي (NCME) وجمعية علم النفس الأمريكية American Psychological Association (APA) كتابًا محدّثًا في العام (2014) عوضًا عن نسخته السابقة التي صدرت في عام (1999) لمعايير القياس التربوي والنفسي بعنوان Standards for Educational and Psychological Testing.

لذا، يتضح أن التقويم عملية ضروريّة يجب أن تؤدّى بعناية فائقة، لذا يتوقع من المعلم أن يكون على علم بكيفية بناء واستخدام أساليب وأدوات التقويم الصفّي المتنوعة خصوصًا مع هذا التغيير الذي يشهده التعليم اليوم (علّام، 2019). كما وتشير بروكهارت (Brookhart, 2013) إلى أن الهدف الأساسي من التقييم هو توفير المعلومات المهمة للمعلمين والطلبة والتي يمكن استخدامها بهدف تطوير التعليم والمخرجات التعليمية.

استجابة لهذه الدعاوي، وفي شهر أكتوبر من العام 2008 تم إطلاق رؤية البحرين الاقتصادية الشاملة (2030) والتي تحدد الطريق أمام استراتيجية التنمية الاقتصادية في البحرين، وأنشأت على إثر ذلك ثلاثة مشاريع رئيسية كجزء من هذه الاستراتيجية وهي: مجلس التنمية الاقتصادية، جامعة بوليتكنك البحرين، وكلية البحرين للمعلمين التي تعنى في العمل على إعداد المعلمين والأخصائيين التربويين، وبذلك تأسست كلية البحرين للمعلمين في عام 2008 بالتنسيق مع وزارة التربية والتعليم وبرامج تعليم المعلمين التي صممها المعهد الوطني للتعليم في سنغافورة (عشر سنوات من كلية البحرين للمعلمين، 2018).

كما تم تأسيس هيئة ضمان جودة التعليم والتدريب في البحرين ضمن مبادرات مشروع تطوير التعليم؛ وهو إحدى المبادرات الرائدة لرؤية البحرين الاقتصادية (2030)، بهدف تحقيق نقلة نوعية وتحسين الخدمات المقدمة في قطاعي التعليم والتدريب المهني في مملكة البحرين، ورفع الكفاءات المهنية، وتحقيق التنمية البشرية لمواطني المملكة (BQA, 2022).

إلا أنه ورغم هذه الجهود فما زال هنالك قصورا في تقييم جودة الممارسات التقييمية للمعلمين في تقييم تحصيل الطلبة، وهذا ما أظهرته العديد من تقارير هيئة جودة التعليم والتدريب في زياراتها للمدارس وتقديمها لتوصيات تتعلق بتقييم مستويات الطلبة. في ضوء ذلك جاءت فكرة هذه الدراسة تحديدا لتوفير مؤشرات حول هذه الممارسات التقييمية في المدارس الحكومية في مملكة البحرين، والنظر بدقة في هذه الممارسات بشتى أشكالها داخل الصف وخارجه وتقييمها وفقًا للمعايير التي أصدرتها مؤسسات التعليم والقياس التربوي العالمية.

الممارسات التقييمية:

تصنف بروكهارت ونتكو (Brookhart & Nitko, 2014) نتاجات التعلّم إلى معرفيّة (Cognitive) تهتم بمهارات التفكير والمعرفة العقلية، ووجدانيّة (Affective) تتعلق بكيف يشعر الطالب، وما الذي يعطيه قيمة، ومهاريّة (Psychomotor) تعنى بالتعامل مع المهارات الحركية والإدراكات الجسدية، ويعدّ تحديد ما يتوقع من الطالب إنجازه خطوة أساسية في العملية التدريسية، حيث يمكن النظر للتدريس من خلال ثلاث ركائز أساسية تعتبر أنشطة متداخلة وهي: تقرير ما يجب على الطالب تعلّمه (تخطيط)، القيام بالتدريس الفعلي (تنفيذ)، تقييم نتاجات التعلّم (تقييم).

وضمن الجهود العالمية لتطوير التعليم والارتقاء بممارسات المعلمين في تقييم الطلبة، أكدت (AFT, NCME, & NEA, 1990) على أن فهم المعلّم لهذه المصطلحات يعد متطلبًا سابقًا للاستخدام الفعال لأساليب التقييم، حيث يستخدم التقييم في تشخيص المشكلات، مراقبة التطور، تقييم التدريس، رصد العلامات، التخطيط للتدريس، تطوير المناهج، وتقييم تطوير المدرسة، كما نبهت إلى ضرورة أن يكون المعلّم قادرًا على استخدام المعلومات الناتجة عن تقييمه في اتخاذ القرارات المتعلّقة بالطلبة.

وقامت هذه المؤسسات بتطوير قائمة بمعايير كفايات المعلمين في مجال التقييم التربوي للطلاب، وتشكلت من سبعة معايير كالتالي:

1- اختيار أساليب تقويم مناسبة للقرارات التعليمية: وتعني قدرة المعلّم ومهارته في اختيار الأساليب المناسبة للقرار التعليمي المعين، كقدرته على استخدام أساليب متعددة كالاختبارات التحريرية والشفوية، وملف الإنجاز، والعروض، وعينات من أعمال الطلبة، والواجبات المنزلية،

وسلالم التقدير، والتقويم الذاتي، والملاحظات، والاستبيانات، والمشروعات وتقويم الأقران.

2- بناء أساليب تقويم مناسبة للقرارات التعليمية: وتعني قدرة المعلّم ومهارته في بناء أدواته لتقييم تحصيل الطلبة وضرورة أن تتميز بالدقة والصدق والعدالة والثبات، وهذا يتطلب من المعلّم أن يكون قادرًا على تحديد جودة أدوات التقويم التي يقوم بإعدادها حيث يعد ذلك أحد أوجه الضعف الرئيسة لدى المعلّمين وقد يرجع ذلك لعدم تدريبهم تدريبًا كافيًا في أثناء تكوينهم في برامج كليات التربية.

3- تطبيق، وتصحيح، وتفسير نتائج أساليب التقويم الداخلي والخارجي: فالاختبارات من الممكن أن يسيء المعلم استخدامها بما يتعارض مع أخلاقيات التقويم، إذ ربما يثير المعلم دافعية البعض من طلابه دون الآخرين، أو يمنحهم وقتًا أطول، أو يهدد البعض بالدرجات، أو غير ذلك مما يتنافى مع مبدأ تكافؤ الفرص. كما أن التصحيح يتأثر بعوامل متعددة تؤثر في النتائج، وعلى المعلّم أن يلتزم بمعايير وتعليمات تصحيح الاختبارات، أما تفسير الدرجات فهو أكثر المهارات صعوبة وتعقيدًا حيث لا يقتصر تفسير الدرجات على عرض الدرجات لأولياء الأمور، وإنما تحتاج لتفسير لاستخلاص معلومات تفيد الطلبة والآباء وتحقق ما يسعون لمعرفته.

4- استخدام نتائج التقويم في اتخاذ قرارات تتعلق بالطلبة، وتخطيط العملية التدريسة، وتطوير المناهج، وتجويد الأداء المدرسي: ويعنى بذلك أن دور المعلم مشارك في اتخاذ القرارات التربوية المؤثرة في مستقبل الطلبة، لذا فمن الضروري أن تكون هذه القرارات صادرة عن ممارسات تقييمية ذات جودة وصادقة قدر الإمكان، حيث يجب أن تستند هذه القرارات لبيانات ومعلومات موثقة، وليس لآراء واعتقادات خاصة.

5- تطوير إجراءات صادقة لوضع التقديرات للطلاب استنادًا إلى تقويمات الطلبة: وتعني مهارة المعلّم في إعداد وتنفيذ وتوضيح الإجراءات التي يستخدمها في رصد علامات الطلبة ووضع تقديراتهم استنادًا إلى التكاليف المختلفة التي تشتمل المشروعات، الأنشطة الصفية، الواجبات المنزلية، الاختبارات الموجزة، الاختبارات التحصيلية. ويتعين على المعلم تبرير مدى منطقية هذه التقديرات وعدالتها، إضافة لقدرته على تقييم وتعديل إجراءات وضع التقديرات هذه بما يسهم في تحسين صدق التفسيرات المستمدة منها.

6- الإخبار بنتائج التقويم للطلاب، والآباء، وغيرهم من المعنيين والمربين: ويعني ذلك قدرة المعلم ومهارته في تقديم التوضيحات المناسبة واستخدام اللغة السليمة عند إبلاغه عن درجات الطلبة غير متأثرٍ بخليفات الطلبة أو ظروفهم الحياتية، ويتوجب على المعلّم أن يأخذ بعين الاعتبار الخطأ في القياس عند استخدام نتائج التقييم في اتخاذ قرارات تتعلق بكل طالب، كما يتعين عليه القدرة على توضيح ما تتضمنه التقارير المكتوبة لنتائج تقييم الطلبة داخل غرفة الصف للمسؤولين في المدرسة وأولياء الأمور.

7- معرفة أساليب التقويم، والاستخدامات غير المناسبة للمعلومات المستمدة من التقويم، والتي تتنافى مع أخلاقيات التقويم، وشرعيته: يتعين على المعلّم اكتساب مهارات فكرية وتطبيقية تتعلق بالقوانين والقرارات المتعلقة بالحالات الخاصة والتي تؤثر في مجال التقويم، كما يجب عليهم إدراك وفهم أن بعض إجراءات التقويم من الممكن أن يُساء استخدامها مما يترتب على ذلك آثار ضارة.

ويؤكد علّام (2019) أن اتفاق هذه الجمعيات العلميّة على هذه المعايير لهو دليل على أهميتها ووظيفيتها للمعلمين، والحاجة لتدريب المعلمين عليها بشكل مستمر قبل توظيفهم أثناء عملهم لاكتساب المهارات المتعلقة بجودة ممارسات المعلمين في تقييم الطلبة وذلك بغية تجويد العملية التعليمية، وإثراء تعلّم الطلبة، وتحقيق المستويات التعلّمية المتوقعة.

وبالرغم من أنَّ هذه المعايير صدرت عن مؤسسات عريقة قدمت إسهامات في المجال التربوي، إلا أن بروكهارت (Brookhart, 2011) قدمت ورقة بحثيّة اعتبرت فيها هذه المعايير قديمة لحدٍّ ما ذلك أنها لم تأخذ بعين الاعتبار المفاهيم الجديدة حول التقويم التكويني كما أنها لم تأخذ بعين الاعتبار كذلك المهارات والمعارف المطلوبة من المعلّم في ظلّ المسؤولية التي تقع عليه في عصر (حركة الإصلاح القائم على المعايير) Standards-based Reform Movement، لذا قام الباحثون تباعًا بتطوير معايير جديدة يجب على المعلّم النظر إليها بعناية عند ممارسته لعملية التقييم التربوي، حيث يجدر الإشارة إلى جهود ماكميلان (McMillan, 2000) الذي حدد أفكارًا كبرى يجب على المعلّم والمدرسة امتلاكها على حدٍّ سواء، وهذه الأفكار هي كالتالي: التقييم يعتبر عملية الحكم بمهنية، يقوم التقييم على مبادئ القياس والتقويم المترابطة، اتخاذ القرارات التقييمية يتأثر بسلسلة من الضغوطات، يؤثر التقييم في في دافعية الطلبة وتعلّمهم، التقييم يحوي أخطاءً، التقييم الجيّد يعزز التدريس والتعلّم، التقييم الجيّد ثمين، التقييم الجيّد عادل وأخلاقي، التقييم الجيّد يستخدم أساليب تقييمية متنوعة، التقييم الجيّد فعال وذو جدوى، والتقييم الجيّد يتضمن التكنولوجيا بشكل مناسب.

ثم توالت الجهود حين برزت جهود ستيغينز (Stiggins, 2009) الذي اقترح خمسة كفايات اعتبرها مفاتيح لجودة التقييم وهي: لماذا نقيّم؟ ويقصد به أن عمليات التقييم ونتائجه تخدم أهدافًا مناسبة وواضحة، وماذا نقيّم؟ ويقصد به أن التقييم يعكس أهداف تعلّم للطلبة واضحة ومهمة، وكيف نقيّم؟ ويقصد به أن يتم ترجمة أهداف التعلّم إلى تقييمات تعطي نتائج دقيقة، وكيف نخبر بالتقييم؟ حيث يتوجب أن تدار النتائج بحرفيّة وأن يتم إيصالها بفاعليّة، وأخيرًا كيف نشرك الطلبة؟ حيث يتم إعطاء الفرصة للطلبة في عملية تقييمهم. بهذا يمكن اعتبار هذه الكفايات الخمس معبرة عن صياغة جيدة لجودة الممارسات التقييمية، كونها قد تجاوزت القصور حول التقويم التكويني ومنحت للطالب مساحة ليكون جزء من عملية التقييم.

وبتظافر جميع هذه الجهود، فقد قامت بروكهارت (Brookhart, 2011) بعمل مراجعة لأغلب ما كتب حول جودة الممارسات التقييمية للمعلّمين،

ثم قامت بتحديد عددٍ من المعايير التي يتعيّن على المعلّم إتقانها في ممارساته التقييمية في القياس والتقويم للطلبة حيث ذكرت أنه يجب على المعلّم أن يكون قادرًا على امتلاك المعرفة التامة بالمحتوى المرتبط بمادة التدريس، والتعبير بوضوح عن النوايا التعليمية التي تتوافق مع محتوى التدريس وعمق التفكير بما تتضمنه أهداف المنهج ومعاييره بطريقة قابلة للقياس والتقويم، وامتلاك ذخيرة مهاريّة من الاستراتيجيات التي من خلالها سيخبر الطلبة كيف سيتم تقييم انجازهم، وفهم أهداف واستخدامات الخيارات التقييمية المتاحة له وأن يكون ماهرًا في استخدامها، وامتلاك المهارة في تحليل الأسئلة الصفّية والفقرات الاختباريّة والمهام المرتبطة بتقييم الأداء، وتقديم تغذية راجعة فعّالة ومفيدة للطلاب حول أدائهم، وبناء معايير تقييمية تساعد في تحديد مستويات أداء الطلبة في الصفوف وتحويلها إلى معلومات ذات قيمة تساعد في اتخاذ القرارات بشأن الطلبة ، والقدرة على الإخبار بوضوح حول نتائج التقييم وتفسيرها ومساعدة الطلبة للاستفادة منها لاتخاذ قرارات تعليمية سليمة، إضافة إلى مسؤولياتهم القانونية والأخلاقية أثناء قيامهم بعملهم.

تقويم تعلّم الطلبة في مملكة البحرين:

أصدرت وزارة التربية والتعليم في مملكة البحرين مؤخرًا (2023) دليلًا استرشاديًا للمعلّمين بهدف توحيد الرؤى والفهم للقياس والتقويم وأدواته وأنواعه، وقد ارتكز هذا الدليل بشكل مباشر على التقويم التكويني لما له من أثر في تحسين جودة المخرجات ومتابعة تحسن الطلبة، وفيما يلي أبرز ما جاء في هذا الدليل مما هو مرتبط بالدراسة الحاليّة.

- فلسفة التقويم في مملكة البحرين

تقوم فلسفة التقويم في البحرين على نظرة تكامليّة بين التقويم والتعليم، حيث أن التقويم يعدّ محسّنًا لعمليات التعليم وضابطًا لنواتجه، فهو ركن رئيس من أركان العملية التعليمية يسهم في تعزيز التعلّم والأداء بما يشمل من معارف ومهارات عملية وتفكيرية وتقنيّة، كما أنه يوفر أدوات للحكم على مدى امتلاك الطلبة لمختلف الكفايات التعليمية، كما أنه يمكن المعلّم من التحقق من مدى توظيف الطالب للكفايات التعليمية في سياقات متعددة وفي حل المشكلات وإنتاج المعرفة، كما أنه يطور قدرة المعلّم في إعداد التقارير وتحديد نقاط الضعف والقوة، وتقديم التغذية الراجعة الفاعلة. (وزارة التربية والتعليم، 2023).

- تقويم تعلّم الكفايات التعليميّة

وتنقسم لثلاثة أنواع، وهي: الكفايات المعرفيّة التي تقيم عبر الاختبارات، والمهاريّة والتي تقيم عبر المهام العمليّة، والوجدانيّة حيث تقيّم من خلال الملاحظة لسلوك الطالب في المواقف التعليمية المختلفة.

الأدوات والأساليب المعتمدة لتقويم تعلّم الطلبة

اعتمد النظام التربوي أساليب تقويمية للتقويم التكويني، كما تم بناء سلالم تقدير لفظية (Rubric) وإعداد مجموعة من الإرشادات والتعليمات لتوجيه المعلّم في أثناء تطبيق كل أسلوب وهذه الأساليب وهي:

1- الملاحظة المنظمة: انتباه مقصود من المعلم للطالب بقصد متابعته في موقف نشط، للحصول على معلومات تفيد في الحكم على أدائه وتقويم مهاراته وقيمه وسلوكه وتوجهاته، ويتكون هذا الأسلوب عند تقييم الطالب من خلاله على مجالين وهما:
2- السلوك الصفي للطالب: وهو كل ما يصدر من الطالب من أفعال أو استجابات وما يعبر عنه من قيم واتجاهات نحو المعلمين أو الزملاء أثناء الموقف التعليمي، ويترجم السلوك الصفي في مدى التزام الطالب بتعليمات الانضباط المدرسي وإرشادات المعلم وبناء علاقة إيجابية قائمة على التقدير والاحترام لعلمه وزملائه.
3- الأداء الأكاديمي: وهو مدى تفاعل الطالب مع العملية التعليمية، عن طريق الاستماع والانتباه الجيد في أثناء الشرح، الحرص على النقاش الإيجابي، التعاون مع زملائه في تنفيذ التكليفات والأنشطة المكلف بها ويتألف أداء الطالب الأكاديمي من الاختبارات والمهام وملف أعمال الطالب.

الدراسات السابقة:

أجرى الدوسري (2004) دراسة هدفت إلى الكشف عن ممارسات المعلمين في التقويم الصفي بالمرحلة الثانوية في البحرين، حيث تكونت عينة البحث من (600) معلم ومعلمة تم اختيارهم عشوائيًا من المدارس الثانوية، واستخدم الباحث استبانة مكونة من (26) فقرة موزعة على 3 محاور وهي (ممارسات المعلم في التقويم الصفي، الأدوات التي يستخدمها المعلم في التقويم، المستويات المعرفية التي يقيسها التقويم)، وأشارت نتائج البحث إلى أن الكثير من المعلمين يستخدمون الأدوات التقليدية كالاختبارات بأنواعها، ويستخدمون عوامل غير مترابطة بالتحصيل بشكل كبير في تقدير درجة طلابهم كالمشاركة الصفية ومجهود الطالب الشخصي، وأوصى الباحث باستخدام الملاحظة المباشرة وإجراء المقابلات مع المعلمين للكشف عن

العوامل الكامنة وراء إصرار المعلمين على استخدام متغيرات غير مرتبطة بتحصيل الطالب.

واجرى قواسمة (2007) دراسة سعت لتقويم ممارسات أعضاء الهيئة التدريسية من وجهة نظر الطلبة في جامعة البحرين والمتعلقة بتقويم التحصيل الأكاديمي للطلبة باستخدام الاختبارات التحصيلية، حيث تألفت عينة الدراسة من (581) طالبًا وطالبة تم اختيارهم بشكل عشوائي من المستويات الدراسية المختلفة في الجامعة، وتكونت أداة الدراسة من (25) فقرة يمثل كل منها إحدى الممارسات التقييمية لتحصيل الطلبة تتعلق بتطبيق الاختبار والوقت والتصحيح وجودة الأسئلة ومراقبة الاختبار، وبعد تحليل البيانات واستخراج المتوسطات وإجراء تحليل التباين الثلاثي وفقًا للمتغيرات (الجنس، المستوى الأكاديمي، الكلية)، فقد أظهرت النتائج قيمًا متواضعة لمستوى الممارسات التقييمية حيث بلغت (64.24)% وأن هناك فروقًا ذات دلالة إحصائية في هذا المستوى لصالح أعضاء هيئة التدريس في كلية التربية بالنسبة لباقي الكليات، ولصالح أعضاء هيئة التدريس الذين يدرسون طلبة السنة الأولى بالنسبة لمن يدرسون باقي المستويات.

كما وقام الشوبكي (2018) بدراسة حول الممارسات التقويمية لدى معلمي المدارس المطبقة لنظام الجودة في الأردن من وجهة نظر القائمين عليها، حيث استخدم الباحث الاستبانة، والمقابلة الشخصية، وتحليل الوثائق الرسمية، وتكونت عينة الدراسة من (742) طالبًا وطالبة، و(393) ولي أمر طالب وطالبة، و(54) منسقا ومنسقة، و(35) معلما ومعلمة، و(16) مديرًا ومديرًا مساعدًا، وأظهرت النتائج أن المعلمين في مجال درجة تطبيق الممارسات التقويمية فقد أظهرت النتائج حول (الهدف من التقويم) أنها فوق المتوسط كاكتشاف أخطاء الطلبة وتصحيحها، والاستفادة من نقاط القوة في تعلم جديد، واقرار الخطط العلاجية لمواطن الضعف. وفيما يتعلق بدرجة تطبيق الممارسات التقويمية (لاستراتيجيات القياس والتقويم التربوي)، وذلك عبر تحليل وثائق خطط الدروس، شيوع التقويم المستند إلى الأداء، يليه التقويم باستخدام الورقة والقلم، كما يشيع استخدام قائمة الرصد وسلم التقدير اللفظي مقارنة ببقية الأدوات. أما فيما يخص (الإمكانات والموارد المتاحة) حسب تحليل نتائج الاستبانة الموجهة للمنسقين والطلبة حول المعلمين فقد أظهر المعلمون درجة مرتفعة بالنسبة للمنسقين بينما كانت الدرجة متوسطة بالنسبة للطلبة. كما أظهرت النتائج بعد استقراء الوثائق المتعلقة (بتحليل وتصحيح وتفسير الاختبارات) والمعدة من قبل المشرفين التربويين أن نسبة كبيرة من المعلمين لا يطبقون تحليل المحتوى عند بناء الاختبار (67.82%)، أما صحة منهجية بناء الاختبار فجاءت مرتفعة، وفيما يخص بناء جدول المواصفات فقد كان متوسط من يطبقونه كافيًا وملائمًا (60%). أما ما يخص (القرارات المستندة إلى نتائج التقويم) فجاءت جميعها مرتفعة من وجهة نظر المدراء ومتوسطة من وجهة نظر أولياء الأمور، وبالنسبة إلى (استخدامات نتائج التقويم) فجاءت جميعها مرتفعة من وجهة نظر المدراء وأولياء الأمور.

كما وأجرى تشينغ ودي لوكا وبراوند ويان وراسلي (Cheng, L., DeLuca, C., Braund, H., Yan, W., & Rasooli, A. 2020) دراسة باستخدام المنهج الوصفي مستهدفين البحث في وجهات نظر المعلمين الصينيين والكنديين لفهم ومعرفة مدى كون التقييم ورصد العلامات ممارسة عابرة للثقافة (Cross-cultural practice) ، حيث تم مقابلة (35) معلمًا كنديًا وصينيًا من المدارس الثانوية لاستكشاف ممارساتهم وقراراتهم التقييمية للطلاب عبر تسع مجموعات تركيز (Focus Group)، ومن خلال التحليل الاستقرائي اتضح أن المعلمين يثمنون وبشدة العدالة والإنصاف في تقييم الطلبة كقيمة مركزية في اتخاذ القرارات حول مستويات طلبتهم، وبينت الدراسة أن العامل المشترك بين المعلمين الكنديين والصينيين هو التحدي المتمثل في الالتزام بممارسات تقييمة عادلة ومنصفة بين الطلبة.

كما تناول جيرسينسون (Gershenson, 2020) كيف يمكن للممارسات التقييمية الصارمة أن تؤثر في تحصيل الطلبة، من خلال اجابته على الأسئلة التالية: كيف يمكن لمعايير الممارسات التقييمية لمعلمي مادة الجبر أن تؤثر في تمكن الطلبة من المادة؟ هل تؤثر معايير الممارسات التقييمية لمعلم مادة الجبر في أداء الطالب في المواد الأخرى المتعلقة بالجبر كالهندسة وفي أدائهم في المراحل المتقدمة قبل التخرج من الثانوية؟ ما خصائص المدرسة والمعلم الذي يظهر جودة في ممارساته التقييمية؟ حيث قام بأخذ بيانات طلبة الصف الثامن والتاسع من المدارس الحكومية لشمال كارولينا على مدار عشر سنوات (2006-2016) وقام بمقارنة درجات الطالب في المادة التي يمنحها له المعلم ودرجة الطالب في اختبار نهاية المقرر (End-of-Course test) وبيّن الباحث أن الطالب الحاصل على درجة عالية في المقرر مقابل درجة منخفضة في اختبار نهاية المقرر فمعلّمه يحقق ضعفًا في الممارسات التقييمية، أما الطالب الحاصل على درجات عادية في المقرر مقابل درجات مرتفعة في اختبار نهاية المقرر فإن معلمه يحقق تميزًا في الممارسات التقييمية، وقام بعد ذلك الباحث بتحليل الدرجات ومقارنتها وتوصل إلى أن الطلبة يتعلمون بشكل أفضل من المعلمين الذين يمتلكون جودة وصرامة ومعايير عالية في ممارساتهم التقييمية، كما أن هؤلاء المعلمين تمكن طلبتهم من تقديم أداء مرتفع في المواد الأخرى المرتبطة بالجبر كالهندسة، إضافة إلى ارتفاع تحصيل هؤلاء الطلبة في جميع المدارس باختلاف خلفياتهم الثقافية والعرقية والاجتماعية ومستوياتهم الاقتصادية، وأظهرت النتائج أن الممارسات التقييمية تبدو أكثر صرامة وشدة وجودة في مدارس الضواحي والمدارس المتوسطة والتي تحوي طلبة بأعداد أكبر.

كما درس كل من الناطور وحجازي والزريقي (AlNatour, Hijazi & AlZauraiqih, 2020) مدى استخدام معلّمي اللّغة الإنجليزيّة لاستراتيجيات وأدوات التقويم الحديث في سياق تدريس اللّغة الإنجليزية لغير الناطقين بها، حيث طبقت أداة الدراسة على (43) معلمًا من منطقة إربد، واتضح من نتائج الدراسة أن مستوى استخدام استراتيجيات وأدوات التقويم الحديث جاءت متدنية، كما لوحظ أن الفقرات التي تقيس مستوى تطبيق المعلّمين

لاستراتيجيات التقويم التقليدي جاءت أعلى من بقية الفقرات، وأوضحت الدراسة أن المعلّمين يميلون لاستخدام استراتيجيات وأدوات التقويم التقليدي، كما اتضح عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha = 0.05$) تعزى لمتغيرات الجنس أو الخبرة أو المؤهل العلمي.

وأجرى إيرغول وسيتين (Ergül & Çetin, 2021) في تركيا دراسة هدفت للتعرف إلى مستويات معرفة المعلّمين للأدب حول القياس والتقويم حيث قاموا بتطبيق أداة الدراسة (Assessment Literacy Inventory) على (189) معلّم من المرحلة الثانويّة من مواد دراسيّة مختلفة، وقد أسفرت نتائج الدراسة عن أن مستويات معرفة المعلّمين كانت منخفضة، كما أن الكفاءة الأعلى للمعلّمين ظهرت في معيار (اختيار طرق القياس والتقويم المناسبة) بينما كان المعيار الأقل هو معيار (تطوير إجراءات صادقة لوضع التقديرات للطلاب استنادًا إلى تقويمات الطلبة)، إضافة إلى وجود فروق ذات دلالة إحصائية في مستويات معرفة المعلّمين تعزى لسنوات الخبرة وللمادة الدراسيّة، وقد أوصت الدراسة بإمكانية إقامة جلسات نقاشية وأنشطة ثقافية وتدريبيّة على الأطر النظرية المتعلّقة بالقياس والتقويم للمواد.

وتناول الخزام (Alkhazam, 2023) تقييم أداء معلمات الرياضيات في المفرق في ضوء ممارساتهن لبعض المعايير المهنيّة للمعلمين من وجهة نظر طالباتهن، حيث تكونت عينة الدراسة من (120) طالبة تم استجبن على أداة الدراسة المكوتة من (54) فقرة ضمن ستة مجالات، تبيّن فيها عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة (α=0.05) في متوسط أداء المعلمات في مجالات أداة الدراسة، بينما أظهرت النتائج وجود فروق دالة إحصائيًا في مستوى أداء المعلمات في مجالي إدارة الصف وتنظيمه والتدريس لصالح المعلمات، بينما لم يتبين وجود فروق دالة إحصائيًا في مستوى أداء المعلمات في كل من مجالات الدراسة الثلاثة: التعامل مع الطلبة وتقويم تعلم الطلبة، والترتيب وتنظيم الإرشادات، بينما أظهرت النتائج وجود فروق دالة إحصائيًا تعزى لأثر الصف الدراسي في جميع المجالات ما عدا إدارة الصف وتنظيمه لصالح الصف الثاني الثانوي، ومجال الأنشطة لصالح الصف الأول الثانوي، بينما أظهرت النتائج وجود فروق دالة إحصائيًا لأثر الفرع التعليمي في جميع مجالات الدراسة لصالح الفرع الأدبي.

وأسفرت نتائج دراسة أبستول وسيرينو وكويفاز وسوميكاد وبيناتل (Apostol et al., 2023) التي طبقت على المعلمين والطلبة عن أن المعلّمين يظهرون تميّزًا ملحوظًا في مهاراتهم في إعداد اختبارات الورقة والقلم، وتقييم الأداء، والتقويم الواقعي، وتحليل الفقرات، وإيصال النتائج للطلبة بفاعلية، كما اتضح أن المعلمين والطلبة على حدٍ سواء كانوا إيجابيين في تصوراتهم حول جميع مجالات الاستبانة، كما اتضح عدم وجود فروق بين الجنسين في مستويات جودة أداء المعلمين في ممارساتهم التقييمية، بينما كانت الفروق واضحةً في مستويات الخبرة مما يدلّ على دور الخبرة في كفاءة المعلم في التقييم.

بهذا يتضح أن الممارسات التقييمية نالت الكثير من الاهتمام في الدراسات السابقة، ويتضح أن هذا الموضوع قد نال مساحة في تنوع مناهج البحث في بعض الدراسات كدراسة الشوبكي(2018) في الأردن الذي مزج بين المنهج الوصفي والمنهج الكمّي عبر الاستبانات وتحليل الوثائق، ودراسة جيرسينسون (Gershenson, 2020) في شمال ولاية كارولينا والذي قام بتحليل درجات الطلبة على الاختبار النهائي للمقرر ودرجة الطالب من قبل المعلّم ومقارنتها لمعرفة أثر مدى صرامة ممارسات المعلمين التقييمية في تحصيل الطلبة.

إلا أنه يبدو وجود قصور واضح في المكتبة العربية البحرينية خصوصًا فيما يتعلق بمعرفة مستوى جودة ممارسات المعلمين في تقييم تحصيل طلبتهم، حيث أن دراسة الدوسري (2004) نظرت فقط للتقييم داخل غرفة الصف ولم تتطرق لبقية الممارسات التقييمية المتعلقة برصد العلامات ومتابعة الواجبات وإدارة الامتحان وإعداده وتفسير نتائجه، كما أنها تعد دراسة قديمة تمت قبل بدء مبادرات مشروع تطوير التعليم في البحرين الذي بدأ بتشكيل هيئة ضمان جودة التعليم والتدريب وتأسيس كلية البحرين للمعلمين (2008)، إضافة إلى أن دراسة قواسمة (2000) كانت موجهة لطلبة الجامعات وليست للمدارس الحكومية، لذا فإن هذه الدراسة ضرورية كونها تختص بالمدارس الحكومية للكشف عن مستويات ممارسة المعلّمين لجودة الممارسات التقييمية خصوصًا في ظل التطور الذي شهدته مملكة البحرين في قطاع جودة التعليم وجودة مخرجات المعلمين وتدريبهم بعد تأسيس كلية البحرين للمعلمين وهيئة ضمان جودة التعليم والتدريب.

مشكلة الدراسة وأسئلتها:

ترتكز عملية تقييم اداء المعلم في مملكة البحرين على أداة واحدة تتمثل في استمارة الملاحظة الصفية الموحدة التي تركز على تقييم أداء المعلم في الموقف الصفي بشكل عام حول استخدامه لاستراتيجيات التدريس والإدارة الصفية ودعم فئات الطلبة المختلفة إضافة الى تناولها بشكل بسيط قدرة المعلم على تقييم أهداف التعلّم، على الرغم من أن تقييم تعلم الطلبة يحتاج مهارات وكفايات عالية الدقة من قبل المعلمين كون الأمر يتعلق بقرارات تتعلق بمصير الطلبة ونجاحهم ورسوبهم. فالمعايير المتوفرة في استمارة الملاحظة لا تتناول بصورة مفصلة معايير الممارسات التقييمية بأنواعها، وإنما تتناول جزئية عامة تتعلق بتحقق بأهداف الموقف التعليمي، فالممارسات التقييمية متنوعة ومتعددة ولا يعد الموقف التعليمي سوى جزء بسيطًا من هذه الممارسات.

وعليه فإن هذه الدراسة ستقوم بتسليط الضوء على الممارسات التقييمية بشكل خاص ومحدد في ضوء معايير الجودة العالمية التي أصدرتها

المؤسسات ذات الشأن التربوي في القياس والتقويم وجودة الممارسات التقييمية من وجهة نظر المعلمين أنفسهم وفق معايير الجودة، وبناء على ذلك يمكن ترجمة مشكلة الدراسة إلى السؤالين التاليين:

1.ما مستوى ممارسة المعلّمين في مملكة البحرين لمعايير جودة الممارسات التقييمية؟

2.هل هناك فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) في مستوى الممارسة لدى المعلمين لمعايير جودة الممارسات التقييمية تعزى لمتغير الجنس أو المؤهل أو الخبرة أو المرحلة التدريسية، أو الخضوع لدورة تدريبية في القياس والتقويم التربوي؟

أهداف الدراسة:

تسعى الدّراسة الحالية لتحقيق الأهداف التّالية:

معرفة مستوى ممارسة المعلمين لمعايير جودة الممارسات التقييمية.

معرفة الفروق في مستوى ممارسة المعلمين لمعايير جودة الممارسات التقييمية.

أهمية الدراسة:

تأتي أهمية هذه الدراسة منبثقة من دور المعلم في تقييم تحصيل الطلبة في المدرسة، حيث أنها تعد مهمة مركزية يتوجب على المعلّم أن يكون على درجة عالية من الإتقان لها، كما يمكن اعتبار هذه الدراسة حاجة ملحّة للتعليم حيث أنها تتعلق بالنظر في الحصول على بيانات حول مستويات ممارسة المعلمين للممارسات التقييمية ذات الجودة والتي ترتبط بتقييم تحصيل الطلبة وذلك وفق معايير صادرة عن مؤسسات عالمية.

فعلى الصعيد العملي، ونظرًا لما تتبناه هذه الدراسة من معايير عالمية لجودة الممارسات التي يتبعها المعلمون، فإنه يمكن الوقوف على نقاط القوة لدى المعلمين، وأبرز نقاط الضعف في جانب الممارسة، كما يمكن ان تسهم في توفير معلومات عن الاحتياجات التدريبية للمعلمين، إضافة لتوفيرها قائمة رصينة من المعايير لدولة البحرين كأداة موثوقة ذات خصائص سيكومترية جيّدة يمكن للمعلّم من خلالها الحكم على مدى جودة ممارساته في تقييم تحصيل الطلبة، وفي تطوير برامج إعداد المعلمين في البكالوريوس والدبلوم العالي للمعلمين.

وأما على الصعيد النظري فسوف تسهم في توفير بيانات في الأدب النظري حول مستويات ممارسة المعلمين في دولة البحرين لمعايير التقييم كما يمكن للباحثين الاستفادة من منهجيتها وأدواتها ونتائجها في بحوث مشابهة.

المصطلحات والمفاهيم النظرية والإجرائية للدراسة:

معايير الجودة Quality Standards:

عرّفت الجمعية الأمريكية للجودة America Society of Quality (ASQ, 2023) بأنها توافر المواصفات أو الخصائص أو الإرشادات التي يمكن للعميل استخدامها باستمرار لضمان ملاءمة هذه الخدمات والمنتجات للغرض منها.

الممارسات التقييمية Assessment Practices:

عرّف هوبا وفيد (Huba & Freed, 2000) الممارسات التقييمية بأنها السلوك المباشر وغير المباشر الذي يقوم به المعلّم في تقييم نتاجات تعلّم الطلبة وتحصيلهم الدراسي، وهي تغطي مجموعة واسعة من القضايا بدءً من معتقدات المعلم وتصوراته حول عملية تقييم تحصيل الطلبة وممارساتهم التي يتقنونها في التخطيط لعملية تقييم التحصيل من مشاركة صفية، وواجبات منزلية، ومشاريع، وتجارب، وبناء الاختبار وتصحيحه ورصد العلامات واستخدام نتائج التقييم، وكتابة التقارير المتعلّقة بذلك.

ويعرف مستوى ممارسة جودة الممارسات التقييمية إجرائيًا بأنه: الدرجة التي يحصل عليها المعلم على قائمة جودة الممارسات التقييمية المستخدمة في هذه الدراسة.

حدود الدراسة ومحدداتها:

تتحدد الدراسة بفئة المعلمين العاملين في المدارس الحكومية في مملكة البحرين للفصل الدراسي الثاني (2023/2022).

حدود الدراسة الزمنية: تمت خلال الفصل الدراسي الثاني للعام الدراسي (2023/2022).

حدود الدراسة الجغرافية: تمت على عينة من المعلمين والمعلمات في محافظات مملكة البحرين.

تتحدد نتائج الدراسة بطريقة جمع البيانات حيث لم يتمكن الباحث من مقابلة جميع المعلمين حضوريًا واقتصرت طريقة الإجابة على الطريقة الإلكترونية.

تتحدد نتائج الدراسة بنوع التدريج الذي استخدم في حساب الدرجة الكلّية حيث كان ثلاثيًا في المعرفة، ورباعيًا في الممارسة.

تقتصر الدراسة على الأداة المستخدمة لجمع البيانات، وهي (قائمة معايير جودة الممارسات التقييمية للمعلمين من إعداد الباحثان) من وجهة نظرهم وما تمتعت به من دلالات الصدق والثبات.

الطريقة والإجراءات:

منهج الدراسة:

تحقيقا لأهداف هذه الدراسة، اتبع المنهج الوصفي التحليلي، وتم جمع البيانات بغرض تحديد مستوى ممارسة المعلمين في المدارس الحكومية لمعايير جودة الممارسات التقييمية.

مجتمع الدراسة:

تكوّن مجتمع الدراسة من جميع معلّمي المدارس الحكومية في مملكة البحرين خلال الفصل الثاني للعام الدراسي (2022-2023)، وقد تعذر على الباحث الحصول على عدد المعلّمين بصورة رسميّة، إلا أنه يمكن تقدير عدد المعلمين بقرابة (14000) معلّم ومعلمة.

عينة الدراسة:

تم اختيار أفراد عينة الدراسة بالطريقة الطبقية العشوائيّة من مدارس مملكة البحرين، وبلغ عددهم (430) معلّمًا ومعلّمة، وذلك كما هو مبين من الجدول رقم (1):

جدول (1): توزيع أفراد العينة من المدارس الحكومية في مملكة البحرين تبعًا لمتغيرات الدراسة

المتغير	المستوى	العدد	النسبة
الجنس	ذكر	204	%47.4
	أنثى	226	%52.6
المرحلة التدريسية	ابتدائية	286	%66.5
	اعدادية	94	%21.9
	ثانوية	50	%11.6
الخبرة	أقل من 5 سنوات	133	%30.9
	من 5 إلى 10 سنوات	102	%23.7
	أكثر من 10 سنوات	195	%45.3
المؤهل العلمي	بكالوريوس	391	%90.1
	دراسات عليا	39	%9.1
الخضوع لدورة تدريبية	لم أتلقَ أي نوع من التدريب أو الدراسة في بناء الاختبارات، أو القياس والتقويم.	74	%17.2
	تلقيت تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى.	181	%42.1
	تلقيت دورة متخصصة فقط في بناء الاختبارات والقياس والتقويم.	68	%15.8
	تلقيت أكثر من دورة متخصصة في بناء الاختبارات والقياس والتقويم.	107	%24.9
	المجموع	430	%100

ح-0ع7غ

6عخمأداة الدراسة:

تحقيقًا لأهداف الدراسة تم تطوير أداة الدراسة قائمة جودة الممارسات التقييمية

أ. خطوات بناء القائمة:

بعد الاطلاع على الأدب النظري والتربوي للقياس والتقويم وما أصدرته مؤسسات القياس والتقويم حول معايير جودة الممارسات التقييمية

(Brookhart, 2011; Stiggins, 2009; McMillan, 2000; AFT, NCME, & NEA, 1990) لاحظ الباحث المعايير التي اتفقت عليها هذه المصادر، حيث أنها تناولت خمسة مجالات رئيسية حددها الباحث كمجالات لقائمة المعايير، وهي تطوير الأدوات والأساليب التقييمية، التصحيح ورصد الدرجات، الإخبار بنتائج التقويم، الممارسات الأخلاقية للتقويم، استخدام نتائج التقويم، لذا قام الباحث بكتابة تجمع الفقرات وإعداد قائمة أولية بمعايير جودة الممارسات التقييمية حيث تألفت القائمة من (56) فقرة موزعة على هذه المجالات الخمسة وبتدريج ليكرت.

وصف القائمة في صورتها النهائية:

بعد كتابة تجمّع الفقرات للقائمة في صورتها الأولية حيث تكوّنت من (56) فقرة وعرضها على التحكيم، تكونت في صورتها النهائية من (53) فقرة، ويتعيّن على المعلّم أن يستجيب للفقرة وفق أسلوب ليكرت وفقًا لتدريج رباعيّ (1-4) بحيث يشير (4= دائمًا)، (3= غالبًا)، (2= أحيانًا)، (1=بتاتًا)، وقد تكوّنت القائمة من خمس مجالات وزّعت عليها الفقرات مرتبةً كالتالي: ممارسات تطوير الأدوات والأساليب التقييمية (26 فقرة) وممارسات تصحيح ورصد الدرجات (5 فقرات) وممارسات الإخبار بنتائج التقويم (7 فقرات) والممارسات الأخلاقية للتقويم (6 فقرات) وممارسات استخدام نتائج التقويم(9 فقرات).

وبهذا أصبحت الأداة جاهزة للتطبيق للتحقق من خصائصها السايكومتريّة من بعد عرضها على محكمين مختصين وتطبيقها على عينة استطلاعية (50) معلّم ومعلّمة ومن خارج العينة.

الخصائص السيكومترية لأداة الدراسة:

ب. الصدق:

صدق المحتوى:

عُرضَت القائمة في صورتها الأوليّة المكونة من (56) فقرة على أساتذة جامعيين في القياس والتقويم وعددهم (5) أساتذة من الجامعة الأردنية وجامعة البحرين وجامعة الخليج العربيّ، كما تم عرضه على معلّمين خبراء وعددهم (3) معلّمين من مختلف المراحل التدريسية ومختلف المواد؛ وذلك للتأكد من مدى ملاءمة الفقرات ووضوحها وسلامتها اللغويّة، ومدى انتماء الفقرة للمجال، وإبداء أي ملاحظات أو تعديلات يرونها مناسبة. وبناءً على آراء المحكمين، تم حذف عدد من الفقرات من القائمة في صورتها الأوليّة إما لتكرارها في مجال آخر أو لدمجها مع فقرة أخرى أو لعدم ملاءمتها للقائمة كالفقرات كما أجرينا عددًا من التعديلات تخصّ الصياغة اللّغويّة للقائمة في صورتها النهائية، وتم الإبقاء على بقيّة الفقرات التي حصلت على إجماع (80%) من آراء المحكمين.

صدق البناء:

حسبت قيم معاملات الارتباط المصحح بين الدرجة على كل فقرة والدرجة الكلية الذي تنتمي إليه الفقرة وكذلك مع الدرجة الكلية على القائمة وذلك للتأكد من الدلالة التميزية، ولكونها مؤشرًا على تحقيق صدق البناء للقائمة بعد تطبيقها على عينة استطلاعية عددها (50) معلّمًا ومعلّمة من المدارس الحكوميّة في مملكة البحرين، كما هو موضح في الجدول رقم (2):

جدول (2): معاملات الارتباط المصحح بين كل فقرة مع الدرجة الكلية للبعد الذي تنتمي إليه والدرجة الكلّية للقائمة (N=50)

رقم الفقرة	معامل الارتباط مع الدرجة الكلّية	معامل الارتباط مع الدرجة الكلّية لمجال الفقرة	رقم الفقرة	معامل الارتباط مع الدرجة الكلّية	معامل الارتباط مع الدرجة الكلّية لمجال الفقرة
1	0.55	0.67	**27**	0.69	0.71
2	0.62	0.69	**28**	0.62	0.73
3	0.63	0.73	**29**	0.61	0.75
4	0.53	0.50	**30**	0.65	0.68
5	0.52	0.48	**31**	0.50	0.61
6	0.61	0.55	**32**	0.66	0.85
7	0.46	0.43	**33**	0.70	0.80
8	0.33	0.32	**34**	0.66	0.79
9	0.42	0.50	**35**	0.64	0.86
10	0.40	0.47	**36**	0.70	0.88
11	0.60	0.69	**37**	0.66	0.73

رقم الفقرة	معامل الارتباط مع الدرجة الكلّية	معامل الارتباط مع الدرجة الكلّية لمجال الفقرة	رقم الفقرة	معامل الارتباط مع الدرجة الكلّية	معامل الارتباط مع الدرجة الكلّية لمجال الفقرة
12	0.65	0.72	**38**	0.72	0.78
13	0.69	0.80	**39**	0.64	0.84
14	0.64	0.72	**40**	0.65	0.87
15	0.65	0.70	**41**	0.64	0.87
16	0.71	0.77	**42**	0.72	0.83
17	0.60	0.69	**43**	0.65	0.80
18	0.68	0.72	**44**	0.66	0.85
19	0.59	0.58	**45**	0.73	0.74
20	0.64	0.68	**46**	0.65	0.71
21	0.72	0.70	**47**	0.75	0.78
22	0.58	0.56	**48**	0.74	0.80
23	0.63	0.63	**49**	0.70	0.77
24	0.70	0.69	**50**	0.69	0.82
25	0.62	0.53	**51**	0.61	0.62
26	0.76	0.73	**52**	0.43	0.61
			53	0.55	0.67

يبين الجدول رقم (2) أن معاملات الارتباط بين درجات فقرات ممارسات تطوير الأدوات والأساليب التقييمية والدرجة الكليّة للمجال تراوحت بين (0.80 و 0.32) وبالدرجة الكليّة للقائمة تراوحت بين (0.76 و 0.33) وتراوحت معاملات الارتباط بين درجات فقرات ممارسات تصحيح ورصد الدرجات والدرجة الكليّة للمجال بين (0.75 و 0.61) وبالدرجة الكليّة للقائمة بين (0.69 و 0.50)، وتفاوتت معاملات الارتباط بين فقرات مجال ممارسات الإخبار بنتائج التقويم والدرجة الكلية للمجال بين (0.88 و 0.73) وبالدرجة الكليّة للقائمة بين (0.72 و 0.64)، وتباينت معاملات الارتباط بين فقرات الممارسات الأخلاقية للتقويم والدرجة الكليّة للمجال بين (0.87 و 0.80)، وبالدرجة الكليّة للقائمة بين (0.66 و 0.64)، أما فيما يخص معاملات الارتباط بين فقرات ممارسات استخدام نتائج التقويم والدرجة الكليّة للمجال فقد تراوحت بين (0.82 و 0.61) وبالدرجة الكليّة للقائمة بين (0.75 و 0.43) وتعد هذه القيم جميعها قيمًا مناسبة تدل على تحقق صدق البناء للقائمة كونها تتفق مع توصيات ثورندايك (Thorndike, 1982) والتي تنص على أن معامل ارتباط الفقرة مع البعد الذي تنتمي له يجب أن يزيد عن (0.30)، كما تم حساب معاملات الارتباط بين المجالات ببعضها وبالدرجة الكليّة للقائمة كما يوضحها الجدول (3).

جدول (3): معاملات الارتباط بين المجالات ببعضها وبالدرجة الكليّة لقائمة الممارسة (N=50)

الممارسة	الدرجة الكليّة	ممارسات تطوير الأدوات والأساليب التقييمية	ممارسات التصحيح ورصد الدرجات	ممارسات الإخبار بنتائج التقويم	الممارسات الأخلاقية للتقويم
تطوير الأدوات والأساليب التقييمية	.947**	1			
التصحيح ورصد الدرجات	.882**	.812**	1		
الإخبار بنتائج التقويم	.835**	.673**	.736**	1	
الأخلاقية للتقويم	.780**	.626**	.635**	.687**	1
استخدام نتائج التقويم	.889**	.751**	.742**	.772**	.745**

** دال عند مستوى (α=0.01)

يتضح من خلال الجدول (3) أن جميع معاملات الارتباط بين المجالات والدرجة الكليّة لقائمة الممارسة كانت موجبة حيث تراوحت بين (0.95 – 0.78) لمجاليّ ممارسات تطوير الأدوات والأساليب التقييمية والممارسات الأخلاقية للتقويم على التوالي، كما أن جميع معاملات الارتباط بين المجالات وبعضها بعضا كانت موجبة حيث كان الارتباط الأعلى بين مجاليّ ممارسات تطوير الأدوات والأساليب التقييمية وممارسات تصحيح ورصد الدرجات (0.83)، والارتباط الأقل كان بين مجاليّ الممارسات الأخلاقيّة للتقويم وممارسات تطوير الأدوات والأساليب التقييمية وبمعامل ارتباط (0.63)، إضافة إلى أن النتائج تشير إلى أن معاملات الارتباط جميعها ذات دلالة إحصائية عند مستوى الدلالة (α=0.01)، وتعدّ قيم هذه الارتباطات مؤشرًا على تحقق صدق البناء لأداة الدراسة كونها أكثر من (0.40) وفقًا لما يفسّره غيلفورد (Guilford) حول تقسيم معاملات الارتباط (Putri, et al., 2019).

ج. الثبات:

تم التأكد من ثبات القائمة بطريقة الاتساق الداخلي عبر استخدام معادلة كرونباخ ألفا كما هو موضح في الجدول (4).

جدول (4): معامل كرونباخ ألفا لأداة الدراسة (N=50)

الممارسة	عدد الفقرات	معامل كرونباخ ألفا
ممارسات تطوير الأدوات والأساليب التقييمية	26	0.938
ممارسات التصحيح ورصد الدرجات	5	0.733
ممارسات الإخبار بنتائج التقويم	7	0.910
الممارسات الأخلاقية للتقويم	6	0.913
ممارسات استخدام نتائج التقويم	9	0.874
الدرجة الكلّية	53	0.968

يبين الجدول رقم (4) أن معامل الثبات المستخرج بطريقة الاتساق الداخلي باستخدام معادلة كرونباخ ألفا لقائمة جودة الممارسات التقييمية لبُعد الممارسة بلغ للقائمة ككل (0.97) ويعدّ هذا الثبات مرتفعًا وفقًا لما يفسّره غيلفورد (Guilford) حول تقسيم معاملات الثبات (Putri, et al., 2019) ويشير لتمتع القائمة بثبات مناسب للاستخدام لأغراض الدراسة. كما أن الثبات لمجالات القائمة قد تراوح بين (0.73 – 0.94) لمجاليّ ممارسات التصحيح ورصد الدرجات وممارسات تطوير الأدوات والأساليب التقييمية على التوالي، وهي قيم مناسبة تدل على ثبات الأداة.

طريقة حساب الدرجة لمستوى ممارسة المعلمين لجودة الممارسات التقييمية:

استخدمت المعادلة (أعلى تدريج – أدنى تدريج) مقسوما على عدد المستويات. وبذلك صنفت مستويات إجابات أفراد العينة لكل فقرة من فقرات بُعد الممارسة على النحو التالي:

- المستوى المنخفض: (1.00 – 2.00)
- المستوى المتوسط: (2.01 – 3.00)
- المستوى المرتفع: (3.01 – 4.00)

متغيرات الدراسة:

1. الجنس وله فئتان: (ذكر وأنثى)
2. المرحلة التدريسية وله ثلاثة مستويات: (ابتدائية واعدادية وثانوية)
3. المؤهل العلمي وله مستويان: (بكالوريوس ودراسات عليا)
4. سنوات الخبرة وله ثلاثة مستويات: (أقل من 5 سنوات، من 5 إلى 10 سنوات، أكثر من 10 سنوات)
5. الخضوع لتدريب في القياس والتقويم وله أربع فئات: (لم أتلقَ أي نوع من التدريب أو الدراسة في بناء الاختبارات، أو القياس والتقويم، تلقيت تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى، تلقيت دورة متخصصة فقط في بناء الاختبارات والقياس والتقويم، تلقيت أكثر من دورة متخصصة في بناء الاختبارات والقياس والتقويم).

المعالجات الإحصائية التي استخدمت في الدراسة:

- استخدمت المتوسطات الحسابية والانحرافات المعيارية للإجابة عن السؤال الأول.
- استخدم تحليل التباين الخماسي (5-Way ANOVA) للإجابة عن السؤال الثاني.

عرض النتائج ومناقشتها:

النتائج المتعلقة بالسؤال الأول: ما مستوى ممارسة المعلمين في مملكة البحرين بمعايير جودة الممارسات التقييمية؟

للإجابة عن هذا السؤال استخرجت المتوسطات الحسابية والانحرافات المعيارية لمستوى ممارسة المعلمين لجودة الممارسات التقييمية، والجدول رقم (5) يوضح هذه النتائج.

الجدول (5): المتوسطات الحسابية والانحرافات المعيارية لمستويات ممارسة المعلّمين لجودة الممارسات التقييمية مرتبة تنازليًا حسب المتوسط الحسابي لكلّ مجال

رقم الفقرة	الفقرات	المتوسط الحسابي	الانحراف المعياري	المستوى
	المجال الأول: ممارسات تطوير الأدوات والأساليب التقييمية.	3.08	0.64	مرتفع
8	وضع الدرجة المخصصة لكل سؤال بشكل واضح أمامه.	3.49	0.82	مرتفع
7	إخراج وتنسيق الاختبار بشكل مرتب من حيث الأسئلة والمساحة المخصصة للإجابة في نفس الصفحة، وترتيب نوعية الأسئلة.	3.40	0.84	مرتفع
6	وضوح الاختبار من حيث الصياغة اللغوية والعلمية.	3.35	0.89	مرتفع
15	مراعاة شروط ومعايير كتابة أسئلة صح وخطأ (True False).	3.30	0.90	مرتفع
9	كتابة الإجابة النموذجية للاختبار قبل تطبيقه على الطلبة.	3.30	0.95	مرتفع
5	كتابة تعليمات الاختبار بشكل واضح.	3.28	0.95	مرتفع
12	مراعاة شروط ومعايير كتابة أسئلة الاختيار من متعدد (Multiple Choice).	3.24	0.93	مرتفع
24	استخدام الأسئلة الشفهية للطلبة.	3.22	0.92	مرتفع
25	استخدام سجل الملاحظات في تدوين (الحضور والغياب، الواجبات، السلوك، المشاركات.. الخ)	3.22	0.91	مرتفع
19	استخدام أدوات متعددة في تقييم أداء الطالب (اختبارات، عروض تقديمية، مشاريع، واجبات، تقارير، ملف إنجاز...)	3.21	0.89	مرتفع
18	التنويع في الفقرات الاختبارية بحيث تقيس جميع مستويات التفكير .	3.19	0.85	مرتفع
16	مراعاة شروط ومعايير كتابة أسئلة الاجابة القصيرة (Short Answer).	3.17	0.97	مرتفع
26	اتساق أسئلة الاختبارات والواجبات والمشاريع مع نتاجات التعلّم.	3.13	0.93	مرتفع
13	مراعاة شروط ومعايير كتابة أسئلة الإكمال (Fill the blank).	3.11	1.01	مرتفع
23	مراعاة الفروق الفردية بين الطلبة عند تقييمهم.	3.09	0.92	مرتفع
21	تقديم الأدلة المختلفة التي تثبت استحقاق الطالب للدرجة التي حصل عليها.	3.08	0.93	مرتفع
14	مراعاة شروط ومعايير كتابة أسئلة المزاوجة (Matching).	3.06	1.03	مرتفع
4	عرض الاختبار على معلّمين آخرين من ذوي الخبرة للتحقق من محتواه.	3.05	0.97	مرتفع
20	استخدام ملف الإنجاز الخاص بالطالب لتقييم تطوره وتقدمه.	3.05	0.92	مرتفع
17	مراعاة شروط ومعايير كتابة أسئلة الإجابة المطولة (Long Answer).	2.98	1.06	متوسط
3	كتابة مختلف أنواع الفقرات لاختبارات الورقة والقلم.	2.95	0.95	متوسط
22	تقييم طلبة الاحتياجات الخاصة بالطرق التي تناسبهم.	2.89	1.03	متوسط
1	تحليل محتوى المقرر عند التخطيط لكتابة فقرات الاختبار.	2.69	1.05	متوسط
2	بناء جدول المواصفات قبل البدء بكتابة فقرات الاختبار.	2.56	1.08	متوسط
10	إعداد سلّم تقدير لفظي (Rubric) للأسئلة المقالية.	2.52	1.14	متوسط
11	إعلان معايير تقييم واضحة باستخدام سلالم التقدير أو قوائم الشطب (Rubric, Checklist, Rating Scale) للمهام المطلوبة من الطالب قبل بدئه بالعمل.	2.49	1.12	متوسط

رقم الفقرة	الفقرات	المتوسط الحسابي	الانحراف المعياري	المستوى
	المجال الثاني: ممارسات التصحيح ورصد الدرجات	3.10	0.75	مرتفع
30	التصحيح بالاستناد الى مفتاح تصحيح واضح.	3.29	0.90	مرتفع
27	اعلام الطلبة مسبقًا حول الأوزان النسبية لما سيتم تقييمهم عليه (المشاركة، الواجبات، المشاريع، الاختبارات القصيرة، الاختبار النهائي والمنتصف، الاختبار العملي..)	3.17	1.00	مرتفع
31	تصحيح أوراق الاختبار دون النظر في اسم الطالب.	3.16	0.91	مرتفع
29	القيام بعملية التصحيح سؤال سؤال بالتتابع .	3.12	1.04	مرتفع
28	السماح للطلبة بمناقشة آرائهم حول توزيع الدرجات.	2.77	1.11	متوسط
	المجال الثالث: ممارسات الإخبار بنتائج التقويم	3.18	0.72	مرتفع
36	إعلام ولي الأمر بمواطن القوة والضعف للطالب.	3.25	0.83	مرتفع
35	إعلام ولي الأمر بمستوى اداء الطالب قبل الاختبار النهائي.	3.24	0.87	مرتفع
32	تقديم تغذية راجعة شفوية أو كتابية للطالب حول مستوى أدائه.	3.21	0.88	مرتفع
34	إيصال نتائج عملية التقييم للجهات المعنية.	3.20	0.95	مرتفع
37	تقديم مقترحات الى ولي الأمر لتحسين مستوى الطالب بناءً على أدائه.	3.13	0.89	مرتفع
33	السماح للطالب بالمراجعة بعد حصوله على التغذية الراجعة.	3.13	0.92	مرتفع
38	تقديم التفسيرات والتبريرات لنتائج تقييم الطلبة للجهات المعنية.	3.08	0.96	مرتفع
	المجال الرابع: الممارسات الأخلاقية للتقويم	3.44	0.62	مرتفع
43	عدم السماح بالممارسات غير الأخلاقية وغير القانونية (كالغش، التلميح، تسريب الأسئلة، تقديم مراجعة مماثلة للأسئلة..).	3.57	0.70	مرتفع
44	تجنب الممارسات غير الأخلاقية وغير القانونية لنتائج التقييم (تهديد الطالب، انتهاك خصوصية..).	3.47	0.87	مرتفع
41	الحفاظ على سرّية الدرجات.	3.45	0.82	مرتفع
39	الاحتفاظ بنسخ ورقية من نتائج الاختبارات.	3.42	0.83	مرتفع
40	الاحتفاظ بنسخ إلكترونية من درجات الاختبارات.	3.40	0.86	مرتفع
42	اتباع الإجراءات الصحيحة عند مراقبة الامتحانات (الالتزام بالوقت، بدون تلميحات، بدون تدخل في الإجابات).	3.37	0.84	مرتفع
	المجال الخامس: ممارسات استخدام نتائج التقويم	3.24	0.70	مرتفع
47	الاستفادة من نتائج التقويم للهدف التعليمي خلال الحصة لتقديم مزيد من التوضيح والشرح.	3.47	0.84	مرتفع
45	الاستفادة من نتائج التقويم القبلي (للمعارف السابقة) عند شرح الدروس.	3.40	0.80	مرتفع
46	الاستفادة من نتائج التقويم التكويني في تقديم الدعم والمساندة.	3.37	0.84	مرتفع
49	الاستفادة من نتائج التقويم في معرفة الأخطاء الشائعة بين الطلبة.	3.34	0.88	مرتفع
48	الاستفادة من المشاركات الصفية للطلبة في دعم تعلمهم.	3.32	0.90	مرتفع
50	الاستفادة من نتائج التقويم في تحديد نقاط تميز وضعف الطلبة.	3.29	0.87	مرتفع
51	استخراج بعض المؤشرات الإحصائية للاختبار (المتوسط، الوسيط، العلامة الأكثر تكرارًا، نسب النجاح والرسوب والإتقان)	3.15	0.98	مرتفع
53	مراجعة أسئلة الاختبار بناءً على تحليل أسئلة الاختبار.	3.00	1.05	متوسط
52	تحليل أسئلة الاختبار من إعداد المعلم بناءً على (معامل الصعوبة + معامل التمييز).	2.82	1.07	متوسط
المتوسط الكلّي لمستوى الممارسة		**3.16**	**0.60**	**مرتفع**

يبين الجدول رقم (5) أن المتوسط الكلّي لمستوى ممارسة المعلّمين لجودة الممارسات التقييمية جاء بدرجة مرتفعة (3.16)، كما كانت متوسطات جميع المجالات بدرجة مرتفعة حيث تراوحت المتوسطات للمجالات بين (3.08 – 3.44)، حيث كان المجال الأقل متوسطًا هو المجال الأول لممارسات تطوير الأدوات والأساليب التقييمية في مستوى ممارسة المعلمين لجودة الممارسات التقييمية وبمتوسط مرتفع بلغ (3.08)، بينما كان المجال الأعلى هو مجال الممارسات الأخلاقية للتقويم وبمتوسط مرتفع بلغ (3.44).

أما فيما يخصّ فقرات القائمة فقد أظهرت النتائج أن متوسط ممارسة المعلّمين لمجال ممارسات تطوير الأدوات والأساليب التقييمية قد تراوحت بين (3.49 – 2.49) وهو المجال الأقل بين المجالات الخمسة لممارسة المعلمين لجودة الممارسات التقييمية، حيث نالت الفقرة (8) لممارسة وضع الدرجة المخصصة لكل سؤال أمامه بشكل واضح أعلى متوسط (3.49) بدرجة مرتفعة ونالت الفقرة (11) لممارسة إعلان معايير تقييم واضحة باستخدام سلالم التقدير أو قوائم الشطب (Rubric, Checklist, Rating Scale) على أقل متوسط (2.49) بدرجة متوسطة. بينما تراوحت متوسطات الفقرات لمجال ممارسات التصحيح ورصد الدرجات بين (3.29 – 2.77) حيث نالت الفقرة (30) لممارسة التصحيح بالاستناد الى مفتاح تصحيح واضح على أعلى متوسط (3.29) بدرجة مرتفعة، ونالت الفقرة (28) لممارسة السماح للطلبة بمناقشة آرائهم حول توزيع الدرجات على أقل متوسط (2.77) بدرجة متوسطة. أما مجال ممارسات الإخبار بنتائج التقويم فقد أظهرت متوسطات ممارسة المعلمين لهذه الممارسات متوسطات مرتفعة لجميع الفقرات تراوحت بين (3.25 – 3.08) حيث إن الفقرة (36) لممارسة إعلام ولي الأمر بمواطن القوة والضعف للطالب نالت أعلى متوسط (3.25) بدرجة مرتفعة، بينما الفقرة (38) لممارسة تقديم التفسيرات والتبريرات لنتائج تقييم الطلبة للجهات المعنية نالت أقل متوسط في المجال (3.08) وبدرجة مرتفعة. وبالمثل كانت فقرات مجال الممارسات الأخلاقية للتقويم حيث أن جميع المتوسطات للفقرات أتت بدرجات مرتفعة تراوحت بين (3.57 – 3.37) وهو المجال الأعلى بين المجالات، حيث حصدت الفقرة (43) لممارسة عدم السماح بالممارسات غير الأخلاقية وغير القانونية (كالغش، التلميح، تسريب الأسئلة، تقديم مراجعة مماثلة للأسئلة..) على أعلى متوسط (3.57) بدرجة مرتفعة، والفقرة (42) لممارسة اتباع الإجراءات الصحيحة عند مراقبة الامتحانات (الالتزام بالوقت، بدون تلميحات، بدون تدخل في الإجابات) على أقل متوسط (3.37) وبدرجة مرتفعة كذلك. أما فقرات المجال الأخير لممارسات استخدام نتائج التقويم فتراوحت المتوسطات بين (3.47 – 2.82) حيث إن الفقرة (47) لممارسة الاستفادة من نتائج التقويم للهدف التعليمي خلال الحصة لتقديم مزيد من التوضيح والشرح نالت أعلى متوسط (3.47) بدرجة مرتفعة، بينما الفقرة (52) لممارسة تحليل أسئلة الاختبار من إعداد المعلم بناءً على (معامل الصعوبة + معامل التمييز) نالت أقل متوسط لهذا المجال (2.82) بدرجة متوسطة.

ويمكن عزو ارتفاع مستوى ممارسة المعلمين لهذه الممارسات إلى القدر الكبير الذي أولته مملكة البحرين لتدريب المعلمين وتمهينهم على جميع ما يخص الاختبارات والقياس والتقويم داخل الغرفة الصفية وخارجها، حيث إنه يتعيّن على المعلّمين اليوم في البحرين ودون استثناء أن يلتحقوا ببرامج كليّة البحرين للمعلمين التي من ضمن مقرراتها الدراسية مقرر يختص بالقياس والتقويم في العملية التدريسية، وطبيعة تدريس هذه المواد في الكلية تستلزم من المعلّمين تطبيق جوانب المعرفة التي يتعلمونها في المادة في سياقات عمليّة، حيث يطلب منهم على سبيل المثال تحليل الاختبارات ونقد صياغة الفقرات وكتابة جداول المواصفات وتحليل المحتوى.

كما يمكن تفسير سبب ارتفاع مستوى ممارسة المعلّمين لجودة الممارسات التقييمية إلى كون المعلّم اليوم أصبح واجبًا من واجباتهم الوظيفية أن يقدم هذا الأداء، حيث تشرف على أدائه بشكل مستمر فرق التحسين المدرسي التابعة للوزارة، إضافة للإشراف التربوي، وإدارة القياس والتقويم، الأمر الذي انعكس على زيادة ممارسة المعلّمين لهذه الممارسات التقييمية.

إضافة إلى أن تأسيس هيئة جودة التعليم والتدريب في العام (2008)، وما قامت به من دور رائد في رفع مستوى جودة أداء المدارس وذلك من خلال الزيارات الدورية التي تقوم بها هذه الهيئة للنظر في الأداء المدرسي بالكامل، حيث تقوم بمراجعة الوثائق وحضور الزيارات الصفية وتحليل الخطط التدريسية للمعلم والخطط التنفيذية والاستراتيجية للمدرسة وللأقسام كما تنظر في جميع جوانب العمل المدرسي، ويعدّ جانب التقييم الذاتي للمدرسة، وتقويم الأهداف، ووضع الخطط العلاجية من أهم الجوانب التي تقوم الهيئة برصدها وملاحظتها والتدقيق عليها؛ لذا فإن المعلمين أصبحوا مطالبين بشدة أن يكونوا على قدرٍ عالٍ من الالتزام بممارسات تقييمية ذات جودة عالية ليرتقوا للنجاح في متطلبات زيارة هيئة جودة التعليم والتدريب.

وتتفق الدراسة الحالية مع ما دعى إليه ماكميلان (McMillan, 2000) حول وصفه للتقييم الجيّد بأنه عادل وأخلاقي حيث جاء في الدراسة الحالية أن المجال الأعلى كان للممارسات الأخلاقية للتقويم، وتتفق مع دراسة (Apostol et al., 2023) والتي أظهرت أن المعلّمين يبدون تميّزًا واضحًا في جميع مجالات التقييم التي طبقوها وبدرجات مرتفعة.

النتائج المتعلقة بالسؤال الثاني: هل هناك فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) في مستوى الممارسة لدى المعلمين لمعايير جودة الممارسات التقييمية تعزى لمتغير الجنس أو المؤهل أو الخبرة أو المرحلة التدريسية، الخضوع لدورة تدريبية في القياس والتقويم التربوي؟

للإجابة عن هذا السؤال تم فحص الفروق بين المتوسطات تبعًا لمتغيرات الدراسة الخمسة باستخدام تحليل التباين الخماسي (5-Way ANOVA) ويوضح الجدول (6) المتوسطات الحسابية والانحرافات المعيارية لمستوى ممارسة المعلمين لجودة الممارسات التقييمية.

الجدول (6): المتوسطات الحسابية والانحرافات المعيارية لمستوى ممارسة المعلّمين لجودة الممارسات التقييمية تبعًا لمتغيرات الدراسة.

المتغير	المستوى	المتوسط الحسابي	الانحراف المعياري
الجنس	ذكر	3.02	0.59
	أنثى	3.06	0.61
المرحلة التدريسية	ابتدائية	3.04	0.61
	اعدادية	3.01	0.61
	ثانوية	3.08	0.54
الخبرة	أقل من 5 سنوات	2.97	0.66
	من 5 إلى 10 سنوات	3.04	0.58
	أكثر من 10 سنوات	3.11	0.56
المؤهل العلمي	بكالوريوس	3.17	0.59
	دراسات عليا	2.91	0.65
الخضوع لتدريب في القياس والتقويم	لم أتلقَ أي نوع من التدريب أو الدراسة في بناء الاختبارات، أو القياس والتقويم.	2.79	0.65
	تلقيت تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى.	3.02	0.58
	تلقيت دورة متخصصة فقط في بناء الاختبارات والقياس والتقويم.	3.15	0.55
	تلقيت أكثر من دورة متخصصة في بناء الاختبارات والقياس والتقويم.	3.21	0.57

يتبيّن من خلال الجدول (6) وجود فروق ظاهرية بين المتوسطات لمستوى ممارسة المعلمين لجودة الممارسات التقييمية، ولفحص دلالة هذه الفروق تم استخدام تحليل التباين الخماسي ويوضح الجدول (7) نتائج تحليل التباين الخماسي.

الجدول (7): تحليل التباين الخماسي لفحص دلالة الفروق بين متوسطات مستوى ممارسة المعلمين بمعايير جودة الممارسات التقييمية.

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	الدلالة الإحصائية
الجنس	0.183	1	0.183	0.542	0.462
المرحلة التدريسية	0.172	2	0.086	0.255	0.775
الخبرة	1.348	2	0.674	1.993	0.138
المؤهل العلمي	2.365	1	2.365	6.992	*0.008
الخضوع لتدريب في القياس والتقويم	8.452	3	2.817	8.329	*0.000
الخطأ	142.060	420	0.338		
المجموع المصحّح	154.685	429			

* ذو دلالة إحصائية عند مستوى الدلالة (α=0.05)

يبيّن الجدول (7) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) في متوسطات مستوى ممارسة المعلمين لجودة الممارسات التقييمية على الأداة ككل وذلك حسب متغيّر المستوى الدراسي ولصالح حملة شهادة البكالوريوس بمتوسط (3.17) مقابل حملة شهادة الدراسات عليا بمتوسط (2.90) كما هو موضح في جدول (16)، كما أظهرت النتائج وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) في متوسطات ممارسة المعلمين لجودة الممارسات التقييمية على الأداة ككل وذلك حسب متغير الخضوع لتدريب في القياس والتقويم، ولمعرفة لصالح من كانت تلك الفروق تم استخدام اختبار (أقل فرق دالّ Least Significant Differenc) للمقارنات البعديّة، والجدول (8) يبين ذلك.

الجدول (8): نتائج اختبار أقل فرق دال (LSD) للمقارنات البعديّة لمتوسطات مستوى ممارسة المعلمين بمعايير جودة الممارسات التقييمية على الأداة ككل وحسب متغير الخضوع لتدريب في القياس والتقويم.

الخضوع لتدريب في القياس والتقويم				
		تلقيت تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى.	تلقيت دورة متخصصة فقط في بناء الاختبارات والقياس والتقويم.	تلقيت أكثر من دورة متخصصة في بناء الاختبارات والقياس والتقويم.
	المتوسط الحسابي	3.018	3.15	3.20
لم أتلقَ أي نوع من التدريب أو الدراسة في بناء الاختبارات، أو القياس والتقويم.	2.79	*0.2287	*0.3720	*0.4244
تلقيت تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى.	3.02		14340.	*0.1957
تلقيت دورة متخصصة فقط في بناء الاختبارات والقياس والتقويم.	3.15			0.0524

* ذو دلالة إحصائية عند مستوى الدلالة (α=0.05)

يتبين من الجدول (8) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) في متوسطات مستوى ممارسة المعلمين لجودة الممارسات التقييمية على الأداة ككل بين من لم يتلقوا أي نوع من التدريب أو الدراسة في بناء الاختبارات أو القياس والتقويم، وجميع الفئات الأخرى التي تلقت تدريبًا أو دراسةً في القياس والتقويم، حيث ظهرت فروقٌ بينهم وبين من تلقوا تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى ولصالح من تلقوا تدريبًا بمتوسط حسابي (3.02) إضافة لمن تلقوا دورة متخصصة في بناء الاختبارات والقياس والتقويم بمتوسط حسابي (3.15) إضافة لمن تلقوا أكثر من دورة متخصصة في بناء الاختبارات والقياس والتقويم بمتوسط حسابي (3.21)، مقابل متوسط حسابي (2.79) لمن لم يتلق نوع من التدريب أو الدراسة في بناء الاختبارات أو القياس والتقويم، كما يتضح وجود فرق دالّ عند مستوى الدلالة (α=0.05) بين المعلمين الذين تلقوا أكثر من دورة متخصصة في بناء الاختبارات والقياس والتقويم والمعلمين الذين تلقوا تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى ولصالح المعلمين الذين تلقوا أكثر من دورة بمتوسط حسابي (3.21) مقابل متوسط حسابي (3.02) لمن تلقوا تدريبًا أو دراسة في بناء الاختبارات والقياس والتقويم ضمن دورة شملت مواضيع أخرى.

ويمكن تفسير الفروق التي ظهرت في متغير المؤهل العلمي ولصالح البكالوريوس، بأن من يحملون درجة الماجستير من المعلّمين لم يأخذوا هذه الدرجة في مجالات التربية؛ وإنما في مجالات تخصصية في المواد التي يقومون بتدريسها كماجستير في الأدب العربي، الفيزياء، الرياضيات، تكنولوجيا التعليم، حيث أن المعلّمين هؤلاء تركز تدريبهم وتعليمهم على المحتوى العلمي للمادة دون التطرق بشكل مكثف لطرق تدريس المادة وإعداد وبناء الاختبارات.

أما فيما يخص تفسير الفروق التي ظهرت في متغير الخضوع لدورة تدريبية في القياس والتقويم، فإنه يتضح تغلّب من أخذ أكثر من دورة مختصة في القياس والتقويم وبناء الاختبارات بشكل ملحوظ؛ ويمكن تفسير ذلك بأن مثل هذه الورش التدريبية والدورات العلمية وبرامج التدريب المهني تكون مركزة وموجهة بشكل واضح ودقيق نحو أساليب القياس والتقويم التي تهمّ المعلّم في العملية التعليمية، كما أنها تمزج بين الطابعين العلمي والعملي؛ فيتوجب على المعلّم أن يبدي براعته العلمية في الأسئلة والأجوبة وحلقات النقاش والعروض التقديمية، وبراعته العمليّة في أداء المهام التي توكل إليه في البرامج التدريبية حول بناء الاختبارات وتصحيحها وتفسير نتائجها وتحليلها وتوثيقها وإيصال نتائجها، كما أن النتائج أظهرت أن من لم يخضع لدورة تدريبية في القياس والتقويم كانوا أصحاب المتوسط الأقل في بعدي المعرفة والممارسة معًا، وهذا لا يدع شكًا أمامنا بأن مستويات المعرفة والممارسة لجودة الممارسات التقييمية لن تأتي مرتفعة من فراغ؛ وإنما يجب أن يكون هناك تدريب مستمر ومتابعة حثيثة لتطوير هذه الممارسات وتحسين أداء المعلمين.

تتفق نتائج الدراسة الحالية مع ما دعت إليه بروكهارت (Brookhart, 2011) حول ضرورة أن يكون لدى المعلمين مستويات عالية من الممارسة لجودة الممارسات التقييمية، وما في ذلك من أثر في تحسين جودة عمليات تقييم الطلبة، كما تتفق جزئيًا مع دراسة أبوستول وآخرون (Apostol et al., 2023) والتي أظهرت عدم وجود دلالة لمتغير الجنس في مستويات ممارسة المعلمين لمعايير جودة الممارسات التقييمية، وتتفق مع دراسة الناطور وآخران (AlNatour et al., 2020) والتي اتضح فيها أن لا أثر للجنس أو الخبرة أو المؤهل العلمي في مستويات ممارسة المعلمين لمعايير جودة الممارسات التقييمية.

توصيات الدراسة:

بناء على نتائج الدراسة؛ فإنه يوصي بما يلي:

- ضرورة إجراء دراسة مماثلة لهذه الدراسة على عينة من معلّمي المدارس الخاصة.
- الاستمرار في تعزيز وتشجيع برامج الدعم التربوي والمهني للمعلّمين لتطوير قدراتهم في تقييم تحصيل الطلبة.
- منح المعلّمين فرصًا أكثر في تدريب زملائهم المعلّمين المستجدين على أبرز الممارسات ذات الجودة.
- إجراء بحث كمي ونوعي يتناول استطلاع الطلبة إضافة لتحليل الوثائق المتعلقة بممارسات التقييم.

المصادر والمراجع

الدوسري، ر. (2004). الكشف عن ممارسات المعلمين في التقويم الصفي بالمرحلة الثانوية. *رسالة الخليج العربي*، 24(90)، 57-89.

الشوبكي، م. (2018). *الممارسات التقويمية لدى معلمي المدارس المطبقة لنظام الجودة في الأردن من وجهة نظر القائمين عليها. أطروحة دكتوراه غير منشورة*، الجامعة الأردنية، عمان، الأردن.

علام ص. (2019). *القياس والتقويم التربوي في العملية التدريسية* (الطبعة السادسة). دار المسيرة.

القواسمة، أ. (2007). تقويم الطلبة لممارسات أعضاء الهيئة التدريسية في تقويمهم لتحصيل طلبتهم في جامعة البحرين. *مجلة كلية التربية*، 23(1)، 90-114.

جامعة البحرين. (2018)، عشر سنوات من كلية البحرين للمعلمين. المنامة، البحرين.

هيئة ضمان جودة التعليم والتدريب BQA. (2022). عن الهيئة: من نحن؟ المنامة، البحرين.

وزارة التربية والتعليم. (2023). دليل المعلّم لتقويم تعلّم الطالب. المنامة، البحرين.

REFERENCES

AlNatour, A., Hijazi, D., & AlZuraiqih, H. (2020). The extent of using modern assessment strategies and tools by teachers of English language in an EFL context. *Jordan Journal of Modern Languages and Literatures*, *12*(2), 205–223. https://doi.org/10.47012/jjmll.12.2.5

Alkhazam, A. M. (2023). Assessing the Performance of Female Mathematics Teachers in Light of Practicing Specific Professional Standards. *Dirasat: Educational Sciences*, 50(2 -S1), 569–580. https://doi.org/10.35516/edu.v50i2 -S1.1498

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association (AFT, NCME, & NEA). (1990). The standards for teacher competence in the educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30-32.

American Society of Quality. (2023). Learn about Quality: What Are Quality Standards? Milwaukee, USA.

Apostol, E. L., Reponte-Sereño, R. R., Cuevas, G. C., Sumicad, R., & Pinatil, M. (2023). Evaluating educators: A comprehensive study of teachers' assessment skills and practices. *Journal of World Englishes and Educational Practices*, *5*(3), 56–69. https://doi.org/10.32996/jweep.2023.5.3.5

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied measurement in Education*, 7(4), 279-301.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: issues and practice*, 30(1), 3-12.

Cheng, L., DeLuca, C., Braund, H., Yan, W., & Rasooli, A. (2020). Teachers' grading decisions and practices across cultures: Exploring the value, consistency, and construction of grades across Canadian and Chinese Secondary Schools. *Studies in Educational Evaluation, 67*, 100928. https://doi.org/10.1016/j.stueduc.2020.100928

Ergül, A. Ö. & Çetin, S. (2021). Measurement and assessment literacy levels of teachers in terms of some variables. *Adiyaman Univesity Journal of Educational Sciences*, 11(1), 26-35.

Gershenson, S. (2020). Great Expectations: The Impact of Rigorous Grading Practices on Student Achievement. Thomas B. Fordham Institute.

Gullickson, A. R. (2005). Student evaluation standards: A paradigm shift for the evaluation of students. Prospects, 35(2), 213-227.

Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning.* Allyn & Bacon, 160 Gould St., Needham Heights, MA 02494.

McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research & Evaluation*, 7(8).

Nitko, A. J. & Brookhart, S. M. (2014). *Educational Assessment of Students* (6th Ed.).

Putri, H. E., Isrokatun, I., Majid, N. W. A., & Ridwan, T. (2019). Spatial sense instrument for prospective elementary school student. *Journal of Physics: Conference Series*, 1318(1). https://doi.org/10.1088/1742-6596/1318/1/012142

Stiggins, R. J. (2009). *Essential formative assessment competencies for teachers and school leaders*. In H. L. Andrade & G. J. Cizek (Eds.), Handbook of formative assessment. New York: Routledge.