الاثار التربوية والأكاديمية والمؤسسية للتفسير غير المتوازن لنتائج تقييم الطلاب لعضو هيئة التدريس في بعض الجامعات العربية (دراسة تحليلية نقدية)

وداد محمد الجودر[١*]

الاستلام : ٦ / فبراير /٢٠٢٦
التحكيم : ١٩ / فبراير /٢٠٢٦
القــبول : ٢٠ / فبراير / ٢٠٢٦

© 2025 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2025 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[١*] كلية التعليم التطبيقي ، جامعة البحرين ، الصخير -مملكة البحرين.
عنوان المراسلة : waljawder@uob.edu.bh

https://doi.org/10.20428/ajqahe.v19i68.3535

الآثار التربوية والأكاديمية والمؤسسية للتفسير غير المتوازن لنتائج تقييم الطلاب لعضو هيئة التدريس في بعض الجامعات العربية

الملخص:

تهدف هذه الدراسة إلى تحليل منظومة تقييم الطلاب لأعضاء هيئة التدريس في بعض الجامعات العربية، مع التركيز على الآثار السلبية الناتجة عن التفسير غير المتوازن لنتائج التقييم من قبل بعض الرؤساء المباشرين، مثل رؤساء الأقسام ومن في حكمهم . وتنطلق من مراجعة منهجية للأدبيات العربية والعالمية حول صلاحية تقييمات الطلاب وحدودها ، مع ربطها بممارسات صفية متكررة في السياقات العربية.

توضح الدراسة أن تقييمات الطلاب قد تعكس بعض عناصر الخبرة التعليمية اليومية، مثل وضوح الشرح وتنظيم المقرر، لكنها تتأثر أيضًا بعوامل لا ترتبط مباشرة بجودة التدريس، مثل توقعات الدرجات، وصعوبة المقرر، وحجم الشعبة، وخصائص المدرس الشخصية والثقافية واللغوية. كما تشير الأدبيات إلى أن العلاقة بين هذه التقييمات ومستوى التعلم الفعلي ليست قوية أو مستقرة بما يكفي لاعتمادها منفردة في قرارات مصيرية.

وتخلص الدراسة إلى أن الخطر الأكبر لا يكمن في التحيزات وحدها ، بل في كيفية تفسير النتائج وتحويلها إلى أحكام وقرارات مهنية قد تظلم عضو هيئة التدريس. كما ترصد أثر ثقافة التقييمات والشكاوى في سلوك الطلاب، مثل التراخي، والمساومة على الدرجات، وتراجع احترام الأستاذ. وتؤكد أن تقييمات الطلاب قد تكون مفيدة إذا استخدمت ضمن منظومة متعددة المصادر، لكنها تصبح ضارة إذا استعملت كمؤشر وحيد ، وتدعو إلى نماذج تقييم أكثر عدالة وفاعلية في الجامعات العربية.

الكلمات المفتاحية: تقييم الطلاب، تقييمات الطلاب أعضاء هيئة التدريس، العدالة المؤسسية، النزاهة الأكاديمية، الجامعات العربية

Educational, academic, and institutional implications of the unbalanced interpretation of student evaluation results for faculty members in some Arab universities

Abstract

This analytical-critical study examines student evaluation of teaching (SET) systems in selected Arab universities, focusing on the consequences of unbalanced interpretations of SET results by some direct supervisors (e.g., department chairs and program directors). Based on a selective systematic review of international and Arab research, the study synthesizes evidence on validity limits and contextual biases affecting SET (e.g., grade expectations, course difficulty, and gender/cultural effects) and links them to recurring classroom practices. The study argues that the most damaging risk is not only bias in student responses, but the administrative use of SET as a sole or near-sole high-stakes indicator. It then outlines key educational and institutional implications, and calls for a balanced, multi-source evaluation framework and clearer policies to prevent misuse of evaluation and complaint mechanisms.

Keywords: SET; faculty performance; institutional justice; academic integrity; Arab universities. Century Gothic

[1*] College of Applied Education, University of Bahrain, Sakhir - Kingdom of Bahrain

*Correspondence address: waljawder@uob.edu.bh

https://doi.org/10.20428/ajqahe.v19i68.3535

الإطار العام للدراسة

المقدمة

شهد التعليم العالي عالميًا توسعًا كبيرًا في الاعتماد على تقييمات الطلاب لأعضاء هيئة التدريس بوصفها أداةً لمراقبة جودة التدريس، واتخاذ قرارات متعلقة بالترقية والتثبيت والحوافز، تحت شعارات المساءلة والشفافية والجودة والاعتماد الأكاديمي، مع افتراض أنَّ الطالب الأقرب لتجربة التعلم اليومية داخل القاعة (Marsh, 1987; Spoored et al., 2013). غير أنَّ مراجعات منهجية ودراسات تجريبية متراكمة - خصوصًا في العقدين الأخيرين - قدَّمت صورة أكثر تعقيدًا؛ إذ تتأثر تقييمات الطلاب بعوامل عديدة لا ترتبط مباشرة بجودة التدريس، مثل سهولة المقرر، وتوقع الطالب لمعدله، وحجم الأعمال المطلوبة، بل وخصائص المدرس المدركة، ما يضعف صلاحيتها حين تستخدم حكمًا وحيدًا على الأداء (Boring, 2017; Crater & Sweet-Cushman, 2022). وفي قياسات التعلم اللاحق أو الاختبارات الموضوعية، تظهر العلاقة بين درجات التقييم وبين التعلم الفعلي ضعيفة أو غير مستقرة في عدد من التحليلات (TTL et al., 2017).

وحذَّرت أدبيات نقدية من أنَّ الاعتماد عالي المخاطر على التقييمات في قرارات العقود والترقيات قد يخلق حوافز معاكسة للتعليم الجيد ، مثل تبسيط المناهج، وتخفيف متطلبات المقرر، وتضخُّم الدرجات بهدف إرضاء الطلاب وتجنب الشكاوى، على حساب العمق العلمي ومخرجات التعلم (Emery et al., 2003; Horn stein, 2017; Strobe, 2020). وتلتقي هذه التحذيرات مع نقاش أوسع حول تحوُّل بعض المؤسسات نحو منطق "الطالب العميل" حيث يتقدم الرضا اللحظي على بناء الكفايات العلمية والانضباط الأكاديمي (Emery et al., 2003) .

وفي السياق العربي، تبنَّت جامعات عديدة استبانات تقييم الطلاب ضمن موجة الجودة والاعتماد ، وغالبًا بوصفها متطلبًا من متطلبات الاعتماد الداخلي والخارجي. وتشير دراسات خليجية وعربية إلى حساسية استجابات الطلاب لعوامل مثل سهولة المقرر، وجنسية الأستاذ، ولغة التدريس، وسمات "الشخصية"، مع محدودية اعتراف بعض الطلاب بهذه العوامل بوصفها تحيزات ينبغي تحييدها (Bad et al., 2006; Alissa & Sulaiman, 2007; Al-Jazari, 2021). لذلك لا تنطلق هذه الدراسة من رفض مبدأ أن يكون للطلاب صوت منظم ، بل من التركيز على لحظة "ترجمة البيانات إلى قرار"؛ أي كيف تقرأ النتائج داخل الأقسام والبرامج؟ وكيف قد تتحوَّل - في بعض البيئات - من أداة تحسين إلى أداة ضغط تعيد تشكيل علاقة الطالب بالأستاذ وثقافة الصف الجامعي.

مشكلة الدراسة

تتمثل المشكلة في الفجوة بين فلسفة السياسة العامة للجودة - التي تؤكد أنَّ تقييمات الطلاب مؤشر ضمن منظومة أوسع - وبين ممارسات تفسيرية انتقائية قد تتعامل مع متوسطات سلبية أو تعليقات فردية بوصفها "أدلة قاطعة" تبنى عليها أحكام عالية المخاطر. ويظهر جوهر المشكلة حين تستخدم النتائج عمليًا للضغط على عضو هيئة التدريس أو لتقويض قراراته الأكاديمية داخل الصف، أو حين يُفهم التقييم ضمنيًا بوصفه أداة عقاب، لا أداة تحسين.

وعليه، تصوغ الدراسة سؤالها المحوري كالآتي: كيف يؤثر التفسير غير المتوازن لنتائج تقييمات الطلاب من قبل بعض الرؤساء المباشرين في تشكيل صورة عضو هيئة التدريس والضغط عليه مهنيًا ، وفي إعادة تشكيل الانضباط والمعايير الأكاديمية داخل الصف الجامعي، مع أن هذه القراءات الفردية لا تمثل بالضرورة الرؤية المؤسسية الكلية ولا الفلسفة المعلنة لأنظمة الجودة والاعتماد؟

أسئلة الدراسة

١. ما الأنماط السلوكية للطلاب غير المنضبطة المرتبطة بثقافة التقييمات والشكاوى في بعض الجامعات العربية؟
٢. ما الآثار التربوية والأكاديمية والمؤسسية للتفسير غير المتوازن لنتائج تقييمات الطلاب؟
٣. هل يمكن اقتراح إطار متوازن لتقييم أداء عضو هيئة التدريس؟

أهداف الدراسة

١. رصد الأنماط السلوكية الطلابية المرتبطة بثقافة التقييمات والشكاوى.
٢. تحليل الآثار التربوية والأكاديمية والمؤسسية للتفسير غير المتوازن لنتائج تقييمات الطلاب.
٣. اقتراح إطار متوازن لتقييم أداء عضو هيئة التدريس في الجامعات العربية.

أهمية الدراسة

تنبع أهمية الدراسة من خمس زوايا مترابطة:

- علميًا/بحثيًا: نقل النقاش من سؤال "ماذا تقيس الأداة؟" إلى سؤال "كيف تقرأ النتائج؟ وكيف تستخدم؟" وربط الأدبيات العالمية بواقع الممارسات الصفية العربية (Spoored et al., 2013; Quash et al., 2024) .
- تربويًا/مهنيًا: إبراز أثر التفسير غير المتوازن على دور الأستاذ كمربٍّ وقائد تعلم، وعلى القلق المهني والاحتراق عندما تتحول التقييمات إلى مساءلة عقابية(ASA, 2020) .
- على مستوى الطالب وثقافة الصف: تفسير كيف يمكن أن تتشكّل لدى بعض الطلاب رسالة ضمنية بأنّ التقييم أو الشكوى وسيلة ضغط إذا غابت التوعية والضوابط.
- مؤسسيًا/إستراتيجيًا: توضيح كيف قد يقود الاعتماد غير النقدي على التقييمات إلى إضعاف المعايير وتضخم الدرجات بما يضر بسمعة المؤسسة ومصداقية مخرجاتها(Strobe, 2020; Horn stein, 2017) .
- سياقيًا عربيًا: التأكيد أنّ الخصوصية الثقافية والتنظيمية للجامعات العربية (حساسية الشكاوى، وتوقعات "الخدمة التعليمية"، وتنوع اللغات والجنسيات) تفرض إطارًا تفسيريًا متوازنًا لا ينسخ النماذج دون تكييف.

الإطار النظري للدراسة

تقييمات الطلاب لأداء التدريس: النشأة والاستخدامات

تطورت تقييمات الطلاب منذ سبعينيات القرن العشرين لتصبح ممارسةً مؤسسيةً شائعةً ضمن ضمان الجودة، وتستخدم عادةً لرصد أبعادٍ من الخبرة التعليمية اليومية، وتزويد المدرس بتغذيةٍ راجعةٍ تساعده على التحسين (Marsh, 1984, 1987; Marsh & Dunking, 1992).وعلى المستوى المؤسسي، تستثمر النتائج في تقارير الجودة، ورصد التحسينات، وأحيانًا في توزيع الحوافز، لكن اتساع استخدامها في قرارات الترقية والعقود يجعل من الضروري تحديد ما الذي تحسن الأداة قياسه؟ وما الذي لا تقيسه؟ ومتى تصبح النتائج بحاجةٍ إلى تفسيرٍ سياقي وإحصائي

الجدل حول الصلاحية: بين التعلم الفعلي والرضا اللحظي

https://doi.org/10.20428/ajqahe.v19i68.3535

تشير تحليلات واسعة إلى أنَّ الارتباط بين تقييمات الطلاب وبين التعلم المقاس موضوعيًا ليس قويًا أو ثابتًا بما يسمح باعتبار التقييم مؤشرًا مباشرًا لجودة التدريس (Clayton, 2009) . وتوضح ميتا-تحليلات أنَّ تقييمات الطلاب قد تعكس رضا الطالب أو راحته أكثر مما تعكس تحسنًا حقيقيًا في التعلم ، وأنَّ بعض قياسات التعلم اللاحق قد تكشف علاقةً ضعيفةً أو حتى عكسية (TTL et al., 2017) . وتضاف إلى ذلك مفارقة معروفة: كلما زادت صرامة المقرر ومتطلباته قد ينخفض الرضا اللحظي لدى بعض الطلاب، بينما قد يكون ذلك مرتبطًا بتعلم أعمق على المدى البعيد. ومن هنا ، تظهر مخاطر مكافأة التدريس عبر متوسطات التقييم وحدها ؛ لأنَّ ذلك قد يدفع - خصوصًا في البيئات عالية الحساسية للشكاوى - نحو تبسيط المناهج، وتخفيف المتطلبات، وتضخم الدرجات (Central, 2003; Strobe, 2020).

تحيزات تقييمات الطلاب

تتداخل في نتائج التقييم تحيزات مرتبطة بالنوع الاجتماعي والعرق/الثقافة واللغة والسمات الشخصية المدركة، إضافة إلى تحيزات تتصل بصعوبة المقرر وتوقعات الدرجات، وحجم الشعبة، ونمط التدريس (Boring, 2017; Fan et al., 2019; Crater & Sweet-Cushman, 2022).

تبرز أهمية التحيزات عندما تستخدم النتائج للمقارنة بين أعضاء هيئة تدريس يدرسون مقرراتٍ مختلفةً أو شُعبًا مختلفة، أو عندما تقرأ المتوسطات دون النظر إلى التباين الداخلي ومعدل الاستجابة والتمثيل الحقيقي للعينة، ودون تحليل نص التعليقات وفق معايير واضحة.

خصوصية الأدبيات العربية والخليجية

تؤكد دراسات خليجية وعربية أنَّ استجابات الطلاب تتأثر بعوامل سياقية، مثل جنسية الأستاذ، ولغة التدريس، وطبيعة المقرر، وتوقعات الدرجات، وأنَّ هذا التأثير قد يكون كبيرًا في بعض البيئات التي تتميز بتعدد الجنسيات واللغات وتفاوت الاستعداد الأكاديمي (Bad et al., 2006; Alissa & Sulaiman, 2007; Al-Jazari, 2021) . كما تشير أدبيات حديثة إلى أنَّ أشكالًا من التحيز قد تظهر حتى في منصات التقييم العامة عبر الإنترنت داخل دول عربية، بما يستدعي يقظةً أكبر عند تفسير النتائج (Lamb et al., 2023) . وتترتب على ذلك حاجة مضاعفة لوجود إرشاداتٍ تشغيليةٍ واضحةٍ لقراءة النتائج، وتدريبٍ للمسؤولين على تفسير البيانات ضمن سياقها، حتى لا تترجم التحيزات إلى ظلمٍ مهني أو قراراتٍ تنظيميةٍ قصيرة النظر.

من صلاحية الأداة إلى ممارسات التفسير الإداري

لا يكفي سؤال "هل الأداة صادقة؟" دون سؤال "كيف تستخدم نتائجها؟". إذ تحذِّر بياناتٌ مهنيةٌ من استخدام (SET)كحكمٍ نهائي في قراراتٍ عالية المخاطر، وتوصي بتعدد مصادر التقييم وفصل التطوير المهني عن المساءلة العقابية .(ASA, 2020; Stark & Freight, 2014) وعندما يتبنَّى بعض الرؤساء المباشرين قراءةً انتقائيةً أو متسرعة للنتائج، ينتقل التحيز من مستوى الاستجابة الطلابية إلى مستوى القرار الإداري، فتتشكل بيئة تدفع التدريس نحو الدفاع عن الذات بدل التركيز على التعلم ، وتنتج تشوهاتٍ في ثقافة الصف والعلاقة بين الطالب والأستاذ.

الدراسات السابقة

يعرض هذا القسم أهم الأعمال التي تناولت تقييمات الطلاب لأداء التدريس (SET) من زوايا ثلاث: (١) دراسات عربية/خليجية تظهر حساسية التقييم لعوامل غير تعليمية، (٢) دراسات دولية تؤكد حدود الصدقية

والتحيزات، (٣) دراسات حديثة تقترح تحسيناتٍ في جمع التغذية الراجعة وتفسير النتائج واستخدامها. وتخدم هذه المراجعة هدفين: تحديد ما استقرَّت عليه الأدبيات بشأن محدودية الاعتماد على التقييمات منفردة، ثم إبراز الفجوة التي تركِّز عليها هذه الدراسة :التحيز لا يظهر فقط في إجابات الطلاب، بل قد يتضاعف أثره عند "التفسير الإداري غير المتوازن" للنتائج.

دراسات عربية وخليجية حول تقييم الطلبة والعوامل المؤثرة فيه

تشير مجموعة من الدراسات في البيئة الخليجية/العربية إلى أنَّ نتائج التقييم تتأثر بعوامل مرتبطة بالمقرر وبالطالب وبالسياق العام أكثر مما يُفترض في الاستخدامات عالية المخاطر. من ذلك، دراسة بدري وعبد الله وكمالي (Bad et al., 2006) في جامعة الإمارات العربية المتحدة، التي اعتمدت تحليلًا كميًا لعددٍ كبيرٍ من استبانات التقييم (نحو ٣١٨٥ استبانة) لاختبار متغيرات غير تعليمية قد تحرف النتائج. وقد بيَّنت النتائج أنَّ خصائص المقرر وخصائص الطلبة ودرجة التساهل في التقدير قد ترتبط بتباين درجات التقييم بمعزل عن جودة التدريس الفعلية، وهو ما يثير سؤال العدالة عند تفسير النتائج تفسيرًا أحاديًا داخل الأقسام.

وفي خطٍ مماثل، تناولت دراسة العيسى وسليمان (Alissa & Sulaiman, 2007) في الجامعة الأمريكية في الشارقة تصورات الطلبة للعوامل التي قد تدخل انحيازًا في التقييم، عبر استبانة على عينة واسعة (٨١٩) طالبًا. وأظهرت النتائج أنَّ توقعات الدرجات وصعوبة المقرر وبعض سمات المدرس قد تؤثر في أحكام الطلبة، ما يعزِّز ضرورة الفصل بين "جودة التدريس" و"رضا الطلبة" عند قراءة المخرجات. وتكشف هذه الدراسة كذلك أنَّ إدراك الطالب لهذه العوامل لا يعني بالضرورة تحييدها أثناء الإجابة؛ وهو جوهر الإشكال عند تحويل البيانات إلى أحكام.

وتفيد بعض الدراسات العربية الحديثة (مثل دراسات الحالة،أو المسوح المحلية) بأنَّ حساسية النتائج لبُنية الأداة وطريقة عرضها للطلبة، وتوقيت تطبيقها قد تكون مرتفعة، وأنَّ الرضا العام قد يظهر دون أن يفسر بالضرورة جودة التعلم أو عمق المخرجات؛ وهو ما يفتح سؤالًا عن "كيف تُقرأ النتائج بعد صدورها؟" أكثر من سؤال "كيف صُممت الأداة؟."

كما تبرز أعمالٌ عربية ركَّزت على الخصائص السيكومترية لأدوات التقييم - مثل البحث الذي تحقق من البناء العاملي والثبات والصدق لأداة تقييم التدريس في جامعة الكويت (جاد الرب والجارالله، ٢٠٢٣) - لتؤكد نقطةً مهمة :جودة الأداة القياسية شرطٌ لازم، لكنه غير كافٍ؛ لأنَّ مرحلة "تفسير النتائج" و"تحديد الدلالات" تظل معرضةً لعوامل إداريةٍ وثقافيةٍ قد تقود إلى قراءاتٍ غير متوازنة.

ويضاف إلى ذلك، ما نبَّهت إليه دراسة حول التحيزات في منصات تقييمٍ إلكترونية في سياقٍ خليجي (Lamb et al., 2023)، حيث تظهر أنماط تقييمٍ مختلفة باختلاف خصائص المُقيِّم وسياق المقرر، بما يعني أنَّ قراءة المتوسطات دون ضبط سياقها أو دون تحييد أثر الاستجابات المتطرفة قد يقود إلى أحكامٍ غير عادلة.

دراسات دولية تناولت تحيزات تقييم الطلبة وحدود صدقيتها

تجمع الأدبيات الدولية على أنَّ تقييمات الطلاب ليست "مرآةً صافية" للجودة، بل مؤشرًا جزئيًا يختلط فيه الرضا بالانطباعات بالعوامل غير التعليمية. وتعدُّ دراسة Boring (2017) من الأعمال التي أظهرت أثر التحيز الجندري ضد النساء في تقييمات الطلبة للتدريس، حتى عند ثبات المحتوى/الأداء، وهو ما يضعف صلاحية استخدامها باعتبارها دليلًا منفردًا في قراراتٍ حساسة مثل التثبيت والترقية.

https://doi.org/10.20428/ajqahe.v19i68.3535

كما يوسّع Crater & Sweet-Cushman (2022) النقاش حول الاستخدامات عالية المخاطر لتقييمات الطلاب، محذرين من تحويلها إلى "عملة إدارية" تختزل فيها جودة الأستاذ ومكانته المهنية، لا سيما مع ما يرافق ذلك من تحيزات مرتبطة بالهوية وبالمقرر وبالتوقعات.

وتأتي التحليلات التي قارنت بين تقييمات الطلاب ومؤشرات التعلم الفعلي لتؤكد أنّ العلاقة ليست دائمًا قوية أو مستقرة بما يبرّر الاستخدام المنفرد) مثل TTL et al., 2017 ، وكذلك Clayton, 2009) . وتدعم هذه الأعمال أطروحة أنّ التقييم قد يقيس جوانب من تجربة التعلم (تنظيم، وضوح، تواصل)، لكنه لا يضمن – بمفرده - قياس "الأثر التعلمي" أو "العمق المعرفي"، خاصة عندما تكون المقارنة بين مدرسين يختلفون في طبيعة المقررات ومستوى صعوبتها.

دراسات حديثة حول تحسين جمع التغذية الراجعة وتفسير النتائج واستخدامها

في السنوات الأخيرة، ركّزت أبحاث جديدة على تحسين تصميم أدوات التغذية الراجعة، وتوقيت جمعها، وطريقة عرضها، بما يساعد على تقليل بعض الانحيازات أو رفع فائدة التعليقات. من ذلك، دراسات تفحص بدائل لأسئلة التقييم التقليدية أو أساليب صياغة البنود لتخفيف الانحياز الجندري، وتحسين قابلية الاستفادة من النتائج (Owen et al., 2024).

وتعرض دراسات أخرى مشكلة "انخفاض المشاركة" في الاستجابات، وكيف ينتج عنها تحيز عدم الاستجابة (Nonresponse Bias)الذي يجعل المتوسطات مضللة عندما لا تمثل العينة الطلبة فعليًا (Papadogiannis et al., 2024). وهنا يبرز بُعدٌ تطبيقي مهم :حتى لو كانت الاستبانة جيدة، فإنّ انخفاض معدلات الاستجابة أو تركُّزها في فئة محددة يضعف صلاحية الاستنتاجات.

وتؤكد أدبيات مقارنة (Pined & Ashour, 2022) أنّ ثقافة المؤسسة (مثل مفهوم "مركزية الطالب") تؤثر في معنى التقييم وفي طريقة تعبير الطالب عن رضاه أو غضبه، ما يوجب إرشادات تشغيلية واضحة لكيفية قراءة النتائج داخل الأقسام وعدم نقلها عبر كليات/جامعات دون ضبط السياق.

خلاصة فجوة الدراسات السابقة

مع تنوع الأدبيات في قياس الصدق والثبات، ورصد العوامل المؤثرة، والتحذير من الاستخدامات عالية المخاطر، إلا أنّ فجوة متكررة تبقى قائمة :قلة الدراسات التي تضع "ممارسات التفسير الإداري" في مركز التحليل :كيف تقرأ النتائج داخل القسم؟ كيف تتحوّل إلى مُساءلة أو ضغط؟ وكيف ترسل رسائل ضمنية تغيّر سلوك الطالب والأستاذ؟ وهذه الفجوة هي ما تسعى هذه الدراسة إلى معالجته، ضمن إطار تحليلي نقدي يربط الدليل البحثي بالممارسات الصفية العربية المتكررة.

إجراءات الدراسة

منهجية الدراسة وطبيعتها

تندرج الدراسة ضمن نمط الدراسات التحليلية النقدية التي تعتمد على مراجعة منتقاة للأدبيات، مع قراءة تفسيرية تربط ما تقوله البحوث العالمية والعربية بما يظهر في الممارسات الصفية في السياق العربي. وقد جرى تتبع دراسات ومراجعات منهجية وميتا-تحليلية تناولت صلاحية تقييمات الطلاب وعلاقتها بالتعلم الفعلي وأنواع التحيزات المتضمنة فيها، إلى جانب تقارير مهنية صادرة عن جمعيات أكاديمية، وهيئات جودة، تحذّر من استخدامها مؤشرًا وحيدًا في قرارات عالية المخاطر) مثل Marsh, 1987؛ Spooren et al., 2013؛ Hornstein, 2017؛ ASA, 2020). Quash et al., 2024؛

ويقوم منطق التحليل على نقطتين:

- أنّ تقييمات الطلاب تنتج "بيانات" ذات حدودٍ منهجيةٍ معروفة (تحيزات، تمثيل، سياق، توقعات).
- أنّ الخطر الأكبر قد ينشأ عندما تترجم هذه البيانات إلى أحكامٍ وقراراتٍ عبر تفسيرٍ انتقائي/متسرعٍ من قبل بعض الرؤساء المباشرين (رؤساء الأقسام، مديري البرامج، المسؤولين الوسطيين).

ومن ثم، لا تقدم الدراسة "حكمًا على جامعةٍ بعينها"، بل تسعى لبناء إطارٍ تفسيري عامٍ يوضح كيف ينتج عن التفسير غير المتوازن سلسلةٌ مترابطةٌ من الآثار على مستوى عضو هيئة التدريس والطلبة والمؤسسة، مع المحافظة على التمييز بين السياسات المؤسسية المعلنة وبين الممارسات الفردية في تفسير النتائج.

حدود الدراسة

هذه الدراسة نظرية تحليلية تستند إلى مراجعة أدبياتٍ وخبراتٍ ميدانيةٍ متراكمة، وليست دراسةً ميدانيةً كميةً ببياناتٍ جديدةٍ من جامعاتٍ محددة؛ لذا، تقدم إطارًا تفسيريًا عامًا أكثر مما تقدم تعميمات إحصائية. كما أنّ تركيزها يتمحور حول السياق العربي بنحوٍ عام.

وتتمثل الحدود الأساسية بالآتي:

- الحدود البشرية: الطلبة بوصفهم مصدر البيانات في الدراسات العربية التي تناولتها المراجعة، ولا سيما المشاركين في تطبيقات استبانات تقييم التدريس في جامعات عربية وردت في الأدبيات.
- الحدود الزمانية: اعتماد الدراسات العربية ضمن فترة (٢٠٠٦/٠١/٠١ إلى ٢٠٢٣/١٢/٣١)، مع الاستئناس بمراجع أجنبيةٍ داعمةٍ حتى (٢٠٢٤/١٢/٠١).
- الحدود المكانية: بعض الجامعات العربية (وفق ما ورد في الدراسات المشمولة بالمراجعة).
- الحدود الموضوعية: منظومة تقييم الطلاب لأعضاء هيئة التدريس، وبنحوٍ أدق: آثار التفسير غير المتوازن لنتائج التقييم عندما يُستخدم بديلًا عن قراءةٍ شاملةٍ متعددة المصادر.

نتائج الدراسة

نتائج السؤال الأول

ما الأنماط السلوكية للطلاب غير المنضبطة المرتبطة بثقافة التقييمات والشكاوى في بعض الجامعات العربية؟

تكشف القراءة التحليلية عن أنماطٍ سلوكيةٍ متكررةٍ ترتبط بثقافة التقييمات والشكاوى، وتنعكس مباشرةً على مناخ الصف الجامعي، وعلى توازن السلطة التربوية داخل القاعة.

أولًا: أنماط تتعلق بالغياب والحضور والالتزام الزمني

يظهر في بعض البيئات الجامعية العربية تعامل بعض الطلاب مع لوائح الحضور والغياب باعتباره خيارًا تفاوضيًا وليس متطلبًا أكاديميًا ثابتًا؛ فيتكرر الغياب دون عذرٍ مقبولٍ طوال الفصل، ثم تظهر مطالبة بإعادة الاختبار أو قبول أعذارٍ متأخرةٍ وضعيفة، مع افتراض ضمني بأنّ من حق الطالب فرصةً إضافيةً مهما كان نمط التزامه. وعندما يلتزم الأستاذ بالسياسة المعلنة منذ بداية الفصل، يُفسَّر ذلك أحيانًا باعتباره تشددًا أو عدم مرونة، ثم يُترجم لاحقًا إلى تقييماتٍ متدنيةٍ أو شكاوى. ويتكرر نمط الدخول المتأخر والخروج أثناء المحاضرة بلا ضرورة، وما يسببه من تشتيتٍ وتسطيحٍ لجودة التعلم، خصوصًا عند غياب دعمٍ مؤسسي واضح لسلطة ضبط الصف.

https://doi.org/10.20428/ajqahe.v19i68.3535

ثانيًا: أنماط داخل المحاضرة تخلّ بالانضباط والاحترام

تظهر ممارسات مثل الاستخدام المتكرر للهاتف المحمول أثناء الشرح للتصفح أو المحادثة أو التصوير مع التنبيه، والحديث الجانبي المستمر الذي يتحوّل إلى ضجيج خلفي، ومقاطعات غير منظمة بأسئلة خارج الموضوع أو تعليقات ساخرة أو تبرم علني من المقرر أو الأستاذ. في هذه الحالات، حين يحاول الأستاذ إعادة ضبط الحوار أو تنظيم التفاعل، قد تسجّل عليه انطباعات من قبيل عدم تقبل الرأي الآخر أو العصبية، وتضخّم أحيانًا في استبانات التقييم.

ثالثًا: أنماط تتعلق بالتقييم والاختبارات والدرجات

تعكس بعض الممارسات سوء فهم للعدالة الأكاديمية: المطالبة بإعادة اختبار مع الغياب دون عذر موثق، أو طلب تغيير موعد الاختبار بدعوى ظروف شخصية غير مثبتة، أو الضغط لتسهيل الامتحان، وكشف نمط الأسئلة، أو طلب بنك أسئلة جاهز. وقد تتطور بعض الحالات إلى مساومة مباشرة على الدرجات باستدعاء الشكوى أو التقييم باعتبارها أداة ضغط. وتؤكد الأدبيات أنّ توقعات الدرجات وسهولة المقرر من أقوى العوامل المؤثرة في تقييمات الطلاب بغض النظر عن جودة التدريس الفعلية (Bad et al., 2006 ، Alissa & Sulaiman, 2007) .

رابعًا: أنماط تمس النزاهة الأكاديمية

تتصل هذه الفئة بظواهر الغش في الاختبارات، والتعاون غير المصرح به في الواجبات الفردية، والاستعانة بآخرين لكتابة المشاريع، وشراء أبحاث جاهزة. وعندما يتعامل الأستاذ مع هذه الأفعال بجدية ويطبق اللوائح، يُعاد تأطيره لدى بعض الطلاب بوصفه متشددًا أو قاسيًا، وتحوّل العقوبة إلى اتهام شخصي للأستاذ بدل ربطها بالخرق الأخلاقي. كما تتكرر الأعذار غير الصادقة للحصول على تمديد أو إعادة اختبار، ومع كل محاولة للتدقيق وطلب وثائق رسمية تتصاعد اتهامات عدم المرونة، في بيئة قد تتحوّل فيها التقييمات إلى أداة انتقامية ضد من يطبق قواعد النزاهة (Hornstein, 2017؛ Quash et al., 2024).

خامسًا: أنماط تتعلق بكرامة المدرس واحترامه

تسجل بعض البيئات أشكالًا من خطاب غير مهني مع أعضاء هيئة التدريس شفهيًا وكتابيًا: نبرة أمرية، رسائل بلا تحية أو تعريف، مراسلات في أوقات غير لائقة، وعبارات ضغط ووعيد ضمني، بل أحيانًا اتهامات للنوايا مثل الادعاء بأنّ الأستاذ يتقصّد طالبًا بعينه أو يتحامل على دفعة كاملة دون دليل موضوعي. ويزداد أثر هذا النمط حين لا يجد الطالب ردعًا مؤسسيًا واضحًا، فيشعر عضو هيئة التدريس بأنّ كرامته المهنية مهددة، وأنّ تطبيقه للوائح قد ينقلب عليه عبر لغة اتهامية في التقييمات والشكاوى (Spooren et al., 2013؛ Cook, 2024).

سادسًا: سوء استخدام التقييمات والشكاوى

بدل أن تكون الاستمارات قناة تغذية راجعة لتحسين التدريس، قد تستخدم أحيانًا أداة انتقام: تخفيض متعمَّد جدًا لدرجات التقييم بسبب جدية الأستاذ أو تطبيقه للنظام، وكتابة تعليقات شخصية لا علاقة لها بجودة الشرح أو تنظيم المقرر. وقد تنسّق حملات تقييم بين مجموعة من الطلاب ضد مدرس مُعين عبر الاتفاق على خفض التقييم، وكتابة ملاحظات متشابهة بهدف الإضرار بسمعته. وعندما تقرأ هذه النتائج إداريًا دون تحليل السياق أو فحص الاستثنائية، تتحوّل الأداة إلى سلاح غير متكافئ يخلّ بميزان القوة داخل الصف.

سابعًا: أنماط تقلل من هيبة المقرر والدور التربوي للأستاذ

يتعامل بعض الطلاب مع المقرر باعتباره مسافًا هامشيًا: حضور متقطع، عدم جلب مواد التعلم، شرودٌ واضح، مذاكرة ضعيفة طوال الفصل، ثم توقع نجاح تلقائي ودرجة جيدة. وتتغذّى هذه الثقافة من ثنائية ضارة: الأستاذ الملتزم باللوائح والمعايير العالية يوصف بأنّه صعبٌ وسيئ، بينما الأستاذ المتساهل الذي يسهّل المتطلبات ويرفع الدرجات يوصف بأنّه طيبٌ وجيد. وترتبط هذه الثنائية بظاهرة تضخم الدرجات، حيث ترتفع المعدلات دون تحسّن موازٍ في الكفايات، عندما يُكافأ التساهل بتقييمات مرتفعة، ويُعاقب الجادون بتقييمات متدنية(Stroebe, 2020 ؛ Hornstein, 2017).

هذه المحاور السبعة توضح أنّ السلوكيات ليست مجرد حالات فردية، بل جزءًا من منظومة ثقافية تتشكل عند نقطة التقاء نظام التقييم المؤسسي، وطريقة تفسير نتائجه إداريًا، ورسائل السياسات الضمنية التي يتلقاها الطالب من أسلوب تعامل الإدارة مع الشكاوى والتقييمات.

نتائج السؤال الثاني

ما الآثار التربوية والأكاديمية والمؤسسية للتفسير غير المتوازن لنتائج تقييمات الطلاب؟

أولًا: الآثار على سلوك أعضاء هيئة التدريس ودورهم التربوي

عندما يشعر عضو هيئة التدريس أنّ بعض الرؤساء المباشرين يتعاملون مع تقييمات الطلاب المنحازة بوصفها حقيقة نهائية، وأنّها قد تستخدم للضغط أو التهديد المهني، يبدأ تعديل سلوكه دفاعيًا. يظهر ذلك في الميل إلى تخفيف الصرامة في ضبط الغياب والتأخر والسلوكيات المخلّة، والتردد في توجيه ملاحظات حازمة للطلاب غير المنضبطين خوفًا من أن تقرأ الحزمية قسوةً أو إساءة معاملة. وقد يمتد السلوك الدفاعي إلى تسهيل الامتحانات، وتخفيف الواجبات، ورفع الدرجات بهدف ضمان متوسطات تقييم مرتفعة، حتى لو جاء ذلك على حساب العمق العلمي ومهام التفكير العليا. وتؤكد الأدبيات أنّ الاعتماد المبالغ فيه على تقييمات الطلاب في قرارات عالية المخاطر قد يدفع إلى تفضيل إرضاء الطلاب على تحديهم تربويًا، بما لا يخدم التعلم الفعلي(Clayton, 2009 ؛ Stroebe, 2020)؛ Hornstein, 2017؛TTL et al., 2017

ثانيًا: الآثار على انضباط الطلاب وثقافة الصف الجامعي

بفعل استجابات بعض الرؤساء المباشرين للتقييمات والشكاوى، قد يتشكل لدى بعض الطلاب انطباعٌ بأنّ السلطة الفعلية في القاعة انتقلت من الأستاذ المتخصص إلى الطالب، وأنّ الشكوى أو التقييم المتدني يمكن استخدامه ضد المدرس. فيزداد تكرار أنماط التأخر والغياب والمطالبة بتسهيلات استثنائية في التقييم، ويصبح الانضباط

https://doi.org/10.20428/ajqahe.v19i68.3535

استثناءً. ويتحوّل التفاعل داخل الصف من شراكة تربوية قائمة على احترام متبادل إلى مساحة تفاوض بين طرف يحمل أداة ضغط، وطرف يحاول تجنب إطلاقها ، وهو ما يقترب من منطق الطالب - العميل في التعليم العالي (Emery et al., 2003).

ثالثًا: الآثار على جودة التعليم والمعايير الأكاديمية

حين يُعاقب الأستاذ الجاد ، ويُكافأ المتساهل عبر قراءة سطحية للتقييمات، تميل المنظومة تدريجيًا إلى تضخم الدرجات، وتبسيط المناهج. تحذف موضوعات عميقة أو صعبة خشية وصف المقرر بالثقل، وترفع الدرجات لتجنب الشكاوى، ويعبر الطالب مقررات عديدة برصيد معرفي ومهاري محدود ، مع درجات مرتفعة، ما يضعف معنى الشهادة ومصداقية مخرجات التعلم. وتؤكد تحليلات ميتا ودراسات نقدية أنّ ارتباط تقييمات الطلاب بالتعلم الفعلي ضعيف/غير مستقر، وأنّ الارتباط الأقوى يكون أحيانًا مع سهولة المقرر لا عمق التعلم (TTL et al., 2017 ؛ Stroebe, 2020)؛Clayton, 2009

رابعًا: الآثار على كرامة المدرس وهيبته وصحته النفسية

يتحول أثر التفسير غير المتوازن من مستوى السياسة إلى قلب التجربة المهنية للمدرس: شعورٌ متكررٌ بالظلم وعدم التقدير حين ترفع تعليقات سلبية مجتزأة باعتبارها دليلًا على مشكلة شخصية، دون اعتبار لسياق المقرر أو سلوك الطلاب أو اللوائح التي طبقت. ومع التكرار ينشأ احتراق وظيفي، وفقدان حماس للتطوير، والاكتفاء بالحد الأدنى لضمان مرور الفصل بسلام ، أو التفكير في مغادرة التدريس أو البحث عن بيئات أكثر إنصافًا. وفي هذه الأجواء، تتآكل هيبة الأستاذ داخل الصف، وتتآكل هيبته في نظر نفسه عندما يُدفع إلى التنازل عن قيمته التربوية لصالح حماية وضعه الإداري (Spooren et al., 2013 ؛Cook, 2024).

خامسًا: الآثار على العدالة المؤسسية وثقافة الجامعة

حين يتلقى عضو هيئة التدريس رسائل ضمنية بأنّ كلمة الطالب هي الحاسمة، وأنّ التقييم المنخفض أو الشكوى الفردية يمكن أن تستثمر ضد الأستاذ دون تحقق كاف، تتآكل الثقة بالمنظومة ويضعف الإحساس بالعدالة. وبالمقابل، يتلقى الطالب رسالة بأنّ المؤسسة قد تتبنّى روايته سريعًا ، وأنّه قادرٌ على الضغط عبر أدوات التقييم والشكاوى. فينشأ تناقضٌ بين خطاب الجامعة حول العدالة وبين ممارسة قد تنحاز لصوت واحد على حساب قراءة متوازنة للأداء، وهو ما تحذّر منه التقارير المهنية التي توصي بربط (SET) بإطار متعدد المصادر وبضوابط تفسير واضحة (ASA, 2020 ؛Merritt, 2014).

سادسًا: الآثار على سمعة الجامعة ومكانتها على المدى البعيد

عندما تتوسع ثقافة مكافأة التسهيل ومعاقبة الصرامة، قد تنتشر صورة مجتمعية عن سهولة النجاح داخل المؤسسة، وأنّ الشكاوى والتقييمات كافية لكسر صرامة الأساتذة. ومع الزمن، قد يلاحظ أرباب العمل ضعفًا في التزام بعض الخريجين ومسؤوليتهم، ويتكون شكٌّ في القيمة الفعلية للشهادة مقارنة بمؤسسات تحافظ على معايير أكثر صرامة. كما أنّ جهات الاعتماد لا تنظر إلى الوثائق فقط، بل تراقب اتساق الخطاب مع الممارسة: هل تحمي المؤسسة لوائح النزاهة؟ هل تقرأ التقييمات سياقيًا؟ أم تسمح لها عمليًا بخفض السقف الأكاديمي؟ (Stark & Freishtat, 2014؛Quash et al., 2024).

سابعًا: الآثار على مفهوم النزاهة الأكاديمية داخل الجامعة

تظهر مفارقة تربوية خطيرة: ترفع الجامعة شعار النزاهة الأكاديمية، ثم تسمح في الممارسة بأنّ تصبح تقييمات الطلاب والشكاوى وسيلة انتقام من الأستاذ الذي يطبّق لوائح الغش والحضور. هنا، يُفرَّغ خطاب النزاهة من مضمونه؛ لأنّ الطالب قد يتعلم أنّ النظام يكافئ المراوغة أكثر مما يكافئ الصدق، إذا رأى أنّ من يضبط الغش أو يرفض أعذارًا غير موثقة يمكن أن يتضرر مهنيًا عبر تقييمات سلبية غير محققة (Hornstein, 2017؛ Merritt, 2014).

مناقشة النتائج:

مناقشة نتائج السؤالين الأول والثاني وتفسير دلالاتها التربوية والمؤسسية:

توضح نتائج الدراسة أنّ التفسير غير المتوازن لتقييمات الطلاب لا يبقى مجرد أرقام في تقرير إداري، بل يتسرب إلى البُنية اليومية للحياة الجامعية: يعيد تشكيل سلوك عضو هيئة التدريس، ويغيّر توقعات الطلاب داخل الصف، ويضغط على المعايير الأكاديمية، ويهزّ إحساس المدرس بكرامته المهنية، ويخلق توترات في العدالة المؤسسية، ويهدد سمعة المؤسسة، ويترك ثغرات في خطاب النزاهة الأكاديمية.

يمكن تكثيف الفكرة المحورية في أنّ التعامل غير النقدي مع تقييمات قد تكون منحازة، خصوصًا عندما تستخدم ضد الأساتذة الجادين والمنضبطين، لا يؤدي فقط إلى تشويه صورة عضو هيئة التدريس والنيل من كرامته المهنية، بل يسهم تدريجيًا في إضعاف الانضباط الطلابي وتخفيض المعايير الأكاديمية وإشاعة ثقافة استهلاكية في التعليم ترى المقرر خدمة قابلة لإعادة التشكيل تحت ضغط التقييم والشكاوى، لا خبرة تربوية مشتركة يتحمل فيها الطرفان مسؤولية متبادلة عن جودة التعلم.

كما توضح المناقشة أنّ الإشكال الجوهري لا يقف عند وجود تحيزات في استجابات الطلاب وحدها، بل يتجاوز إلى نقطتين متلازمتين: تصميم منظومة تقييم أداء عضو هيئة التدريس ككل، ثم الوعي التربوي والإداري لدى الرؤساء المباشرين الذين يقرؤون النتائج ويترجمونها إلى قرارات أو إشارات رسمية. حتى لو عُدِّلت الاستبانة، وخُفّفت بعض مظاهر التحيز إحصائيًا، يبقى الخطر قائمًا إن لم تتغير طريقة دمج هذه الأداة في منظومة أوسع، وإن لم تضبط السلطة التأويلية الممنوحة لمن يقرأ النتائج، ويقيم عليها قرارات أو مواقف.

وتؤكد الأدبيات أنّ تقييمات الطلاب لا تقيس التعلم الفعلي بصورة مستقرة، وأنّها ترتبط بدرجة أكبر بالرضا اللحظي وعوامل مثل سهولة المقرر، وتوقعات الدرجات، وخصائص الأستاذ (Boring, 2017؛ Bad et al., 2006؛ Alissa & Sulaiman, 2007؛ TTL et al., 2017). هذا لا يعني أنّ الأداة بلا قيمة، بل يعني أنّ قيمتها مشروطة بإطار الاستخدام: يمكن أن تكون نافعة للتغذية الراجعة والتحسين إذا وُضعت في حجمها الطبيعي ولم ترفع إلى مرتبة الحكم الوحيد.

وتشير النتائج كذلك، إلى أنّ جزءًا من الخلل في بعض الممارسات العربية لا يكمن في وجود التقييم، بل في انتقاله من كونه مؤشرًا مساعدًا ضمن حزمة مؤشرات إلى كونه مرآة شبه وحيدة، تحاكم بها صورة المدرس. وهذا الانتقال قد لا يحدث بقرار مركزي مكتوب، بل يتشكل عمليًا عبر ممارسات بعض الرؤساء المباشرين الذين يفضلون رقمًا واضحًا وسهل التداول (متوسط التقييم) على قراءة متكاملة لملف المدرس تشمل عينات أعمال الطلاب، ومحتوى المقرر، وتغذية راجعة من النظراء، ومؤشرات تعلم موضوعية.

https://doi.org/10.20428/ajqahe.v19i68.3535

من هنا، تدفعنا هذه الدراسة إلى توسيع سؤال الأدبيات التقليدي: ليس فقط هل الأداة صادقة وثابتة؟ بل كيف تستخدم فعليًا في سياقاتنا؟ من يملك سلطة تفسيرها؟ وبأي منطق تقرأ الأرقام والتعليقات؟ هذا الانتقال من نقد الأداة إلى نقد المنظومة التي توظف الأداة ينسجم مع الدعوات الحديثة التي تؤكد أنّ إصلاح التقييم لا يتحقق بتحسين صيغ الأسئلة فقط، بل بمراجعة البُنية الثقافية والإدارية المحيطة بالتقييم(Hornstein, 2017 ؛ Quash et al., 2024).

وأخيرًا، تبرز المناقشة ثلاث خلاصاتٍ مركزية، تمهد للقسم التالي من البحث:

الخلاصة الأولى: أنّ تحيزات تقييمات الطلاب حقيقة موثقة، ولا يجوز تجاهلها عند تفسير الأرقام.
الخلاصة الثانية: أنّ الخطر الأكبر لا يكمن في وجود الأداة، بل في تمكين قراءاتٍ فرديةٍ وانتقائيةٍ لها من قبل بعض الرؤساء المباشرين بما يحولها إلى وسيلة ضغطٍ تقوّض الدور التربوي والأخلاقي لعضو هيئة التدريس.
الخلاصة الثالثة: أن الإصلاح يمرُّ عبر إعادة هندسة منظومة تقييم الأداء لتستوعب تقييمات الطلاب ضمن إطارٍ متعدد المصادر، وتضبط ممارسات التفسير، وتخضع استخدام النتائج لسياساتٍ مكتوبةٍ واضحةٍ تحمي الأطراف، وتخدم رسالة الجامعة المعرفية والتربوية معًا

.

مناقشة السؤال الثالث:

هل يمكن اقتراح إطارٍ متوازنٍ لتقييم أداء عضو هيئة التدريس في الجامعات العربية؟

تقترح الدراسة إطارًا متوازنًا يمكن أن تسترشد به الجامعات عند صياغة سياساتها في تقييم أداء عضو هيئة التدريس، يقوم على مبادئ عمليةٍ تضمن الاستفادة من صوت الطالب دون تحويله إلى حكمٍ أحادي أو أداة ضغط، وبما يحمي الدور التربوي للأستاذ ويصون العدالة المؤسسية. ويستند الإطار إلى ما أكدته الأدبيات من ضرورة تقليل الاستخدام عالي المخاطر لتقييمات الطلاب منفردة، والاتجاه إلى منظومةٍ متعددة المصادر، مع ضوابط تفسيرٍ واضحة(Hornstein, 2017 ؛ ASA, 2020).

تعدد مصادر التقييم

لا ينبغي أن تستخدم استبانات تقييم الطلاب منفردةً في أي قرار عالي المخاطر (ترقية/تثبيت/جزاء/حرمان من مكافأة)، بل تدمج مع مصادر أخرى تكمل الصورة، وتقلّل التحيز. ومن أبرز مكونات الملف المتكامل:

- مراجعة نظراء منظمة (Peer Review) وفق نموذج واضح ومعايير معلنة.
- أدلة مخرجات تعلم الطلاب: نماذج من أعمالهم، جودة المشاريع، مستويات الأداء وفق(rubrics) ، أمثلة من أسئلة الاختبارات ومدى اتساقها مع مخرجات التعلم.
- ملف تدريسي :(Teaching Portfolio) يقدمه عضو هيئة التدريس موثقًا لتصميم المقرر وتطويره، واستراتيجيات التدريس، وخطط التحسين المبنية على التغذية الراجعة.
- تقييم إشرافي/لجنة متخصصة تربط الدليل بالمعايير، وتقرأ سياق المقرر والشعبة.

بهذا، يتحوّل التقييم من حكمٍ أحادي البُعد إلى ملفٍ مهني يوازن بين الخبرة الصفية وبين الدليل التعليمي القابل للتحقق(Hornstein, 2017 ؛ Onwuegbuzie, Daniel, & Collins, 2007).

ضبط ممارسات التفسير لدى الرؤساء المباشرين

تؤكد الدراسة أنَ نقطة الخلل الحاسمة ليست وجود الاستبانة فقط، بل السلطة التأويلية في قراءة نتائجها. لذا، يقترح الإطار ما يلي:

- تدريب منهجي لرؤساء الأقسام/البرامج على تفسير نتائج التقييم في ضوء تحيزاتها وحدودها، بدل قراءتها رقمًا مجردًا.
- عدم بناء الحكم على فصل واحد أو شعبة واحدة، واعتماد الاتجاه العام عبر فصول عديدة، مع مراعاة خصائص المقرر (صعوبة/مستوى/متطلبات/لغة التدريس/حجم الشعبة).
- تحليل التوزيع والانحراف والتطرف، وليس المتوسط وحده، وقراءة معدل الاستجابة وتمثيل العينة قبل أي استنتاج.
- توثيق النقاش مع عضو هيئة التدريس حول النتائج قبل رفع توصية، وإظهار كيف جرت الموازنة بين صوت الطالب وبين الأدلة الأخرى

سياسات مكتوبة تحمي من الاستخدام التعسفي للتقييمات

يقترح الإطار تثبيت مبادئ الحماية في اللوائح، حتى لا تترك العدالة للاجتهادات الفردية، وذلك عبر:

- نصٍّ صريح بأنَّ تقييمات الطلاب مؤشر تغذية راجعة، وليست دليلًا قاطعًا بمفردها.
- فصل الاستخدام التطويري عن الاستخدام المساءلة قدر الإمكان: ما يُستخدم لتحسين التدريس لا يُستخدم تلقائيًا للعقاب.
- أي قرار عالي المخاطر يعتمد على تقييمات الطلاب يجب أن يمرَّ بآليات تحقق وتظلم واضحة، تتيح للأستاذ تقديم سياق المقرر ووثائقه) خطة المقرر، rubrics، أمثلة أعمال الطلاب، تطبيق لوائح الحضور والنزاهة.
- وضع معيار للتعامل مع التعليقات المسيئة أو الشخصية: تستبعد لغويًا وأخلاقيًا من التداول الإداري، ولا تحتسب ضمن الحكم المهني.

إعادة تعريف دور التقييم لدى الطلاب

حتى يكون صوت الطالب مفيدًا، لا بد أن يُفهم التقييم باعتباره مسؤوليةً أخلاقيةً لا منصةً للانطباعات الشخصية. لذا، يقترح الإطار:

- برامج توعية للطلاب تشرح هدف التقييم وحدوده، وأنَّ التقييم مسؤولية نزاهة وموضوعية، وأنَّ الاتهامات الشخصية أو اللغة المسيئة غير مقبولة.
- إدماج رسائل التوعية في تهيئة الطلاب الجدد، وحملاتٍ دوريةً داخل الجامعة، وربطها بثقافة النزاهة الأكاديمية.
- توضيح الفرق للطالب بين "الصرامة الأكاديمية" وبين "سوء التدريس"، وبين "عدالة التقويم" وبين "سهولة الحصول على درجة".

تضمين الإطار في وثائق الجودة والاعتماد

https://doi.org/10.20428/ajqahe.v19i68.3535

يقترح البحث أن تقدِّم الجامعات هذا التصور في وثائق الجودة بوصفه نموذجًا يوازن بين الاستماع لصوت الطالب وبين حماية الدور التربوي للأستاذ، بحيث تنسجم الممارسة مع الخطاب الرسمي عن النزاهة والإنصاف. فالاعتراف المؤسسي بأنَّ قراءة التقييم تحتاج خبرة تربوية وتحليلًا سياقيًا، لا حسابًا رقميًا بسيطًا، يزيد قدرة الجامعة على تحقيق متطلبات الاعتماد دون الإضرار برسالتها التعليمية.

استنتاجات الدراسة

أظهرت الدراسة أنَّ الاعتماد غير النقدي على تقييمات الطلاب في السياق العربي لا يقتصر أثره على مستوى الأفراد، بل يمتد إلى بُنية الثقافة الجامعية. فمن جهة يدفع الأستاذ إلى تبنِّي سلوك دفاعي يقدِّم حماية الذات المهنية على الأهداف التربوية العميقة، ومن جهة أخرى، يعيد تشكيل توقعات الطلاب وسلوكياتهم بحيث تضعف ثقافة الانضباط والمسؤولية المشتركة داخل الصف.

كما بيَّنت الدراسة أنَّ هذه المنظومة تترك آثارًا بعيدة المدى على جودة المخرجات التعليمية، وكرامة المدرس وهيبته، وعدالة المؤسسة، وسمعة الجامعة، ومفهوم النزاهة الأكاديمية. وبالاستناد إلى الأدبيات العالمية والإقليمية، يتضح أنَّ تقييمات الطلاب يمكن أن تكون أداةً مفيدةً للتغذية الراجعة وتحسين الممارسات، لكنها تصبح خطرًا على العدالة التربوية عندما تستخدم بوصفها مؤشرًا وحيدًا أو حاسمًا في القرارات عالية المخاطر (ASA, 2020؛ Hornstein, 2017).

وتنتهي الدراسة إلى ضرورة إعادة وضع تقييمات الطلاب في حجمها الطبيعي باعتبارها أداةً ضمن منظومة تقييم متوازنة، تعتمد على تعدُّد المصادر، وضبط ممارسات التفسير لدى الرؤساء المباشرين، وحماية كرامة عضو هيئة التدريس، وتوعية الطلاب بمسؤولياتهم الأخلاقية في التقييم. وهي بذلك لا تقف ضد الجودة ولا ضد صوت الطالب، بل تدعو إلى عدالة أكبر في توزيع الأدوار والمسؤوليات داخل منظومة التقييم الجامعية.

توصيات الدراسة

- إجراء دراسات نوعية تستكشف بعمق خبرات أعضاء هيئة التدريس مع التقييمات والشكاوى، وكيف تدار إداريًا داخل الأقسام.
- إجراء دراسات كمية تعتمد على تحليل سجلات تقييمات فعلية، وربطها بمتغيرات مثل نوع المقرر، مستوى الطالب، سياسات الغياب والتصحيح، ومؤشرات تعلم لاحقة.
- تطوير نماذج بديلة لتقييم التدريس في الجامعات العربية (متعددة المصادر)، وتصميمها واختبارها تجريبيًا من حيث عدالتها وفاعليتها.
- إجراء دراسات مقارنة بين دول عربية مختلفة، وبين الجامعات الحكومية والخاصة لاستكشاف الفروق في ثقافة التقييم والانضباط.

Publication Ethics Declaration

Conflicts of Interest

The author declares that there are no conflicts of interest. —Financial, personal, or institutional—Related to the preparation or publication of this manuscript.

Funding Information

This research received no external funding.

Authors' Contribution Statement

"The sole author completed all tasks related to the study's conception, writing, and revision".

المراجع

Abrams, P. C. (2001). Improving judgments about teaching effectiveness using teacher rating forms. *New Directions for Institutional Research, 2001*(109), 59–87. https://doi.org/10.1002/ir.4

Alissa, A., & Sulaiman, H. (2007). Student evaluations of teaching: Perceptions and biasing factors. *Quality Assurance in Education, 15*(3), 302–317. https://doi.org/10.1108/09684880710773183

Al-Jazari, F. S. (2021). The potential in student evaluation of teaching for EFL improvement: Voices from the Gulf. *Cogent Education, 8*(1), 1888670. https://doi.org/10.1080/2331186X.2021.1888670

American Sociological Association. (2020). *Statement on student evaluations of teaching*.

Bad, M. A., Abdulla, M., Kamal, M., & Doreen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. *International Journal of Educational Management, 20*(1), 43–59. https://doi.org/10.1108/09513540610639585

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics, 145*, 27–41. https://doi.org/10.1016/j.jpubeco.2016.11.006

Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education, 44*(5), 495–518. https://doi.org/10.1023/A:1025492407758

Clayton, D. E. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education, 31*(1), 16–30. https://doi.org/10.1177/0273475308324086

Cook, S., Watson, D., & Webb, R. (2024). Performance evaluation in teaching: Dissecting student evaluations in higher education. *Studies in Educational Evaluation, 81*, 101342. https://doi.org/10.1016/j.stueduc.2024.101342

Emery, C. R., Kramer, T. R., & Than, R. G. (2003). Return to academic standards: A critique of student evaluations of teaching effectiveness. *Quality Assurance in Education, 11*(1), 37–46. https://doi.org/10.1108/09684880310462027

Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R., & Johnston, E. L. (2019). Gender and cultural bias in student evaluations: Why representation matters. *PLoS ONE, 14*(2), e0209749. https://doi.org/10.1371/journal.pone.0209749

Hornstein, H. A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education, 4*(1), 1304016. https://doi.org/10.1080/2331186X.2017.1304016

Lamb, J. N., Aleissa, H., & Telha, A. (2023). Are student evaluations of teaching influenced by the gender and nationality of instructors? Evidence from the Middle East. *Assessment & Evaluation in Higher Education*.

Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases. *Journal of Educational Psychology, 76*(5), 707–754.

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(3), 253–388.

Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective. In J. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 8).

Merritt, D. J. (2014). Bias, the brain, and student evaluations of teaching. *St. John's Law Review, 82*, 235–287.

Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. T. (2007). A meta-validation model for assessing the validity of student teacher evaluations. *Quality & Quantity, 41*, 353–380.

Owen, J., et al. (2024). Reducing bias in student feedback instruments: Evidence and design principles. *Higher Education Research & Development*.

Papadogiannis, G., et al. (2024). Nonresponse bias in student evaluations of teaching: Implications for interpretation. *Assessment & Evaluation in Higher Education*.

Pined, C., & Ashour, S. (2022). Student evaluations of teaching in the Arab Gulf: Student-centeredness and the politics of quality. *Higher Education Research & Development*. https://doi.org/10.1080/07294360.2022.2156483

Quash, R., et al. (2024). Rethinking student evaluations of teaching in higher education: A systematic review. *Educational Research Review*.

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research, 83*(4), 598–642.

Stroebe, W. (2020). Student evaluations of teaching encourages poor teaching and contributes to grade inflation: A theoretical and empirical analysis. *Basic and Applied Social Psychology, 42*(4), 276–294.

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation, 54*, 22–42.

Wen, D., Sun, X., & Yan, D. (2022). The quality movement: Where are we going? Past, present and future. *Total Quality Management & Business Excellence, 33*(1–2), 92–112. https://doi.org/10.1080/14783363.2020.1801342

Yirdaw, A. (2016). Quality of education in private higher institutions in Ethiopia: The role of governance. *SAGE Open, 6*(1). https://doi.org/10.1177/2158244015624950

جاد الرب، هـ. ف.، والجار الله، ر. س. (٢٠٢٣). البناء العاملي والخصائص السيكومترية لاستبانة تقييم الأداء التدريسي لدى هيئة التدريس بجامعة الكويت. *المجلة التربوية، ٣٧*(146)١٣، –45. https://doi.org/10.34120/joe.v37i146.3729