

Book Review: *Developing Evaluative Judgement in Higher Education: Assessment for Knowing and Producing Quality Work*

In the current landscape marked by the emergence of Generative Artificial Intelligence (GenAI), our educational methods and assessment strategies are confronting unprecedented challenges. A prevailing worry centers on the possible erosion of a vital educational endeavor – the cultivation of evaluative judgment. This concern stands out as one of the most significant issues in contemporary education, as the insufficiency of students' 'capability to make decisions about the quality of work of self and others' (Tai et al., 2018, p.471). To elucidate the concept of evaluative judgment, this volume, edited by Boud and colleagues, stands as the inaugural work exclusively centered on the theory and practices pertaining to this subject. The book is structured into four sections, including 'conceptualising evaluative judgement', 'alternative theoretical perspectives on evaluative judgement', 'approaches to developing evaluative judgement', and 'evaluative judgement for practice and work'.

Section 1 begins with a chapter by Ajjawi, Tai, Dawson, and Boud, delving into the significance, characteristics, historical evolution, and strategies for cultivating evaluative judgement. In the contemporary landscape, higher education confronts a pressing demand to provide graduates with the ability to learn and adapt amidst unprecedented rates of change. Proficient evaluative judgement is imperative for decision-making within the professional sphere and for effective functioning. In contrast to self-assessment as a confined activity, evaluative judgement is distinguished by its capacity to transcend individual perspectives and immediate tasks. While this perspective may seem innovative, its roots trace back to Sadler's (1989) exploration of the role of formative assessment in shaping students' 'evaluative knowledge' and 'evaluative expertise'. Adopting a social constructivist standpoint, additional strategies for nurturing evaluative judgement encompass the use of rubrics, exemplars, self-assessment, peer assessment, and feedback. The application of these strategies is expounded upon in Sections 3 and 4 of this book. In Chapter 2, Dall'Alba explores issues pertaining to the cultivation of evaluative judgement in a digitally enabled world, considering both epistemological and ontological dimensions. The epistemological dimension centers on the enhancement of students' knowledge and competence, whereas the ontological dimension focuses on their development as human beings. The amalgamation of these dimensions in enhancing evaluative judgement in a digitally enabled world entails the utilization of technologies for pedagogical objectives. This includes assessing progress towards professional aspirations, fostering an informed stance in the learning journey, encouraging creativity and diversity in professional development, and mitigating bias

in evaluative judgement. Adopting a generative standpoint, Goodyear and Markauskaite assert in the third chapter that evaluative judgement can be viewed as an epistemic capability, valuable for assessing one's proficiency in engaging in knowledgeable action within specific, dynamically changing situations. They illustrate how a diverse array of examples in professional knowledge work can be classified within a taxonomy of epistemic games, including a detailed examination of a particularly relevant type: the evaluation game. Their insights propose that by amalgamating their perspectives to formulate representations of how epistemic work is accomplished in their respective fields, for instance, by identifying sets of epistemic games, staff members can enhance their capacity to design developmental pathways for the progressive development of students' evaluative judgement. Additionally, such an approach aids in cultivating students' ability to integrate diverse perspectives when forming such judgments.

In Section 2, contributors expound upon alternative theoretical viewpoints regarding the conceptualization of evaluative judgement. These encompass the sociomaterial perspective, historical outlook, the dual-process approach, and the learning process in digital environments. In Chapter 4, Ajjawi and Bearman delve into an exploration of standards from a sociomaterial perspective. The term 'reified' is employed to convey the transformation of standards from mere concepts into tangible artefacts, exemplified by the likes of rubrics or competency frameworks. These standards, in the form of tangible objects, wield the capacity to both propel and confine human actions. The disjunction between the artefact and its interpretation is reconciled by conceptualising a standard as an activity in which individuals actively engage. Thus, a standard assumes a dual nature: a steady, unchanging artefact and an unpredictable, mutable interpretation. The cultivation of students' evaluative judgement serves to empower them, enabling the enactment of standards in a consistent manner for the given task. This is achieved through supportive measures, encompassing dialogues and practical exercises, that assist students in comprehending the concept of quality across various tasks and, significantly, across different institutions. In Chapter 5, Nelson adopts a historical lens to delineate three distinct types of evaluative judgement. The first, termed 'hard evaluative judgement', involves discrimination based on truth. The second, 'soft evaluative judgement', entails discrimination grounded in the concept of importance. The third, 'dynamic evaluative judgement', centers around discrimination based on opportunities. Each type necessitates a

reconciliation with persistent uncertainty, with only the first category capable of retrospective validation through proof, while the categories of soft and dynamic evaluative judgement, richly imbued with cultural contingencies, rarely find retrospective validation, given that an alternative course of action might have led to an unknown and potentially superior outcome. The advancement of the second and third forms of evaluative judgement faces impediments within the framework of constructive alignment. To enhance students' discriminatory agility, there is a call for a learning culture that extends beyond the confines of predefined learning outcomes, encouraging exploration and engagement outside the rigid educational grid. In the subsequent chapter, Joughin highlights the dual processing systems inherent in human judgement. One is an automatic, effortless, unconscious, and intuitive process classified as System 1. The other is a conscious, deliberate, effortful, analytical, and rational approach designated as System 2. The dual system theory, along with the specific heuristics and biases associated with it, makes a significant contribution to our understanding of student evaluative judgement. The key to nurturing students' capacity for evaluative judgement lies in the growing awareness of the benefits associated with System 2, coupled with an acknowledgment of the perils stemming from an unwitting dependence on System 1. In Chapter 7, Lodge, Kennedy, and Hattie highlight the suitability of digital environments for both investigating and intervening during the production of student work. The ongoing application of fundamental psychological research, encompassing self-evaluation of learning and illusions of competence, to the realm of learning in digital environments, is generating opportunities for real-time interventions aimed at enhancing student learning processes and strategies.

In Section 3, the book's emphasis undergoes a transition towards practical methods for cultivating evaluative judgement. This shift includes two chapters that delve into self-regulated learning, along with others that explore computer-generated exemplars and feedback, the analysis of exemplars, the formulation of technology-enabled dialogic feedback, the application of assessment management, and the utilization of degree-level learning outcomes. In Chapter 8, Panadero and Broadbent advocate for a self-regulated learning perspective, delineating four distinct levels (observation, emulation, self-control, self-regulated) in the context of evaluative judgement. They endorse a developmental approach to the latter, shedding light on how students can progressively acquire it. The primary implications of this approach suggest that students must engage in deliberate practice to master a new skill. This practice is optimally enhanced by observing models and receiving feedback as it unfolds. In Chapter 9, aiming to underscore the concept of metacognitive monitoring, Bennett, Lockyer, Kennedy, and Dalgarno outline certain principles shared by models of self-regulated learning and practices of developing evaluative judgement. This effort aims to bridge the divides between parallel research traditions. They offer a concise case study illustrating a well-intentioned, student-centered approach to designing an open-ended online task, providing students with latitude

to interpret the task requirements and seek support from the social context. However, the distributed approach to teaching and formative feedback poses challenges to students' ability to make effective judgments. In Chapter 10, Dawson scrutinizes insights garnered from computer-based evaluative judgement and explores the potential for these lessons to contribute to the advancement of student evaluative judgement. The focus is on discerning lessons learned and avoiding pitfalls, aiming to establish connections between empirical insights derived from artificial intelligence and educational theories about evaluative judgement. The chapter concludes by delineating three principles: first, exemplars should be drawn from authentic real-world practices, showcasing diversity and equality; second, exemplars should be scaffolded with annotations, signals, and summaries; third, exemplars should be employed in a dialogue, emphasizing feedback and explainability. In Chapter 11, Carless, Chan, To, Lo, and Barrett highlight that the theoretical foundation for employing exemplars is grounded in the concept of tacit knowledge, elements that are challenging to convey verbally or in writing. An essential objective of analyzing exemplars with students is to assist them in initiating the acquisition of tacit knowledge by rendering visible some of the expert thinking and judgments of the teacher. They also view the integration of exemplars into the curriculum as a pedagogic tool for cultivating student capacities in evaluative judgement. Additionally, it serves as a means of elucidating assessment requirements through scaffolded use and fostering productive dialogues. In Chapter 12, Henderson, Phillips, and Ryan outline a framework for dialogic feedback processes related to assessment, facilitated by both multimodal and social media platforms. Within this design, these media are utilized to enable and amplify multiple feedback loops, both before and after assessment submission. Within these loops, educators and students articulate their evaluative thinking, fostering a process of reflection and dialogue. The subsequent two chapters delve into the practical aspects of cultivating evaluative judgement within programmes, depicting iterative cycles of activities over time. These chapters hold specific relevance for educational developers and policymakers. In Chapter 13, Ellis emphasizes the crucial need for the learning context, along with the utilization of electronic assessment management and assessment analytics, to be structured in a manner that fosters and promotes nested, iterative opportunities for the development of evaluative judgement. It is imperative for educators and institutions to reconsider the conventional understanding of assessment and feedback, spanning from teaching and learning activities at the unit level to the broader scope of the entire programme. In Chapter 14, Thompson and Lawson highlight three strategies concerning the amalgamation of curriculum design and assessment, centered on holistic degree outcomes, to provide an educational environment fostering evaluative judgement. This integration is achieved through clear, consistent communication to all stakeholders, accompanied by comprehensive data on a student's performance as they progress. These strategies pivot the assessment focus from mere certification towards a more developmental role, aiding both students and staff in

comprehending the objectives they aim to accomplish at the degree level.

The last section proposes that evaluative judgement is not generic but, instead, draws upon and serves as a means of introducing students to disciplinary methods and values. Three chapters within this section draw upon research conducted with learners in healthcare education. In Chapter 15, Bearman proposes that making evaluative judgments in higher education is intricately linked to the academic discipline of the assessment. Consequently, rendering an evaluative judgement involves the discipline's surface structures (or external markers of quality, such as might be encapsulated in a rubric), deep structures (tacit understandings about what quality might signify within the discipline), and implicit structures (how the notion of quality reflects the core values of the discipline). However, students also bring an agentic approach to understanding disciplinary quality, considered a positive facilitator of independent professional practice, which could be encouraged with numerous practical methods, for instance, co-production of criteria and teachers' role-modelling reflexivity. In Chapter 16, Tai and Sevenhuysen propose that peer-assisted learning (PAL) is a fundamental process through which students' evaluative judgement can be cultivated in various learning situations, spanning from the classroom to practical settings. They illustrate this connection through two case studies featuring medical and physiotherapy students: informal PAL arises spontaneously in the presence of a peer, providing opportunities for sharing ideas and practical skills, while formal PAL involves specific teaching activities assigned to the pair or student group. In Chapter 17, Johnson and Molloy focus on the potential for learners to develop evaluative judgement during feedback conversations in healthcare, offering an illustrative example that involves comparing self-assessment with the educator's judgement. They also present recommendations to improve opportunities for evaluative judgement through feedback conversations. In Chapter 18, Rees, Bullock, Mattick, and Monrouxe present the revelations and concealments within junior doctors' narratives regarding preparedness for practice, specifically in terms of evaluative judgement. Moreover, they advocate for exploring the potential of longitudinal audio-diaries as a means to gain insight into the development of trainees' evaluative judgments over time. In the concluding chapter, Boud, Dawson, Tai, and Ajjawi offer practical considerations for redesigning curricula to foster evaluative judgement. From a learner-centric approach, these considerations encompass the discernment of quality, understanding judgement processes, managing biases, assessing the trustworthiness of sources, and actively seeking opportunities for practice. Among these, the most crucial condition is that all students must transition from viewing themselves as consumers of courses, where responsibility lies with others, to becoming active learners. They need to recognize that they are the primary agents for their own learning, and only through their initiative can they fully benefit from available opportunities.

Through emphasizing how students can reap advantages from discerning the quality of their own work and that of their peers, this book establishes a shared discourse that underpins

a substantial portion of preceding research on theoretical and disciplinary perspectives. Following the conceptualization of evaluative judgement in higher education, there is a growing body of empirical research explicitly situating itself within this innovative framework. Recent relevant literature has contested certain insights from these chapters and made fresh contributions to the advancement of evaluative judgement research, traversing from conceptualization to empirical research, thus laying the groundwork for future inquiries.

To begin with, Luo and Chan (2023a) asserted that evaluative judgement competence is not a fixed trait that a student possesses or lacks, but rather a process through which students navigate multiple interconnected facets. This not only advocates for a more 'process-oriented' approach to cultivating students' evaluative judgement skills in the curriculum but also underscores the necessity for an integrated curriculum design that addresses the various aspects emphasized by evaluative judgement and the synergies between them.

Secondly, regarding the multidimensional construction of evaluative judgment, researchers have undertaken further exploration through qualitative and quantitative studies. Building upon three rounds of Delphi surveys involving 14 international experts, Luo and Chan (2023b) conceptualized evaluative judgement as a multidimensional construct encompassing knowledge, attitudes, competencies, actions, and identity-related aspects. This multidimensional perspective aids in dissecting evaluative judgments concerning the development of integrated competencies into readily discernible objectives for teachers and students, thereby enhancing the likelihood of accomplishing these objectives within the curriculum. On this foundation, Luo et al. (2023) fashioned and validated a self-report instrument designed to gauge engineering students' evaluative judgments of intercultural competence. A total of 815 Chinese engineering students partook in the pilot and formal validation phases. The final questionnaire comprises 27 items covering six dimensions. The study furnishes a valuable tool for educators and policymakers, and the instrument facilitates students' self-reflection, assisting them in attaining greater autonomy in guiding their own development and growth.

Thirdly, concerning the practices related to cultivating evaluative judgment development proposed in this book, Ibarra-Sáiz et al. (2020) introduced a predictive model for students' competence development based on peer assessment practices. They illustrated the relationships between the variables of evaluative judgement, participation, feedback, self-regulation, and assessment quality. By scrutinizing four years of data from an undergraduate project management class, this paper formulates a causal model validated using the PLS-SEM method. It elucidates the relationship between these variables and considers the impact of student competence, as well as the mediating nature of feedback and self-regulation in this process.

Concludingly, it is noteworthy that several chapters in this book focus on the utilization of technology in digital environments to enhance students' evaluative judgments. Notably, Generative AI-powered conversational interfaces such as ChatGPT possess the potential to revolutionize the

realms of teaching, learning, and assessment. Therefore, it is timely to engage in a discussion concerning the responses of teachers, and educational institutions to GenAI to explore its current impact on assessment practices and contemplate how it might shape the landscape of evaluative judgement in the future.

REFERENCES

- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2020). Developing student competence through peer assessment: The role of feedback, self-regulation and evaluative judgement. *Higher Education, 80*, 137-156. <https://doi.org/10.1007/s10734-019-00469-2>
- Luo, J., & Chan, C. K. Y. (2023a). Exploring the process of evaluative judgement: The case of engineering students judging intercultural competence. *Assessment & Evaluation in Higher Education, 48*(7), 951-965. <https://doi.org/10.1080/02602938.2022.2155611>
- Luo, J., & Chan, C. K. Y. (2023b). Conceptualising evaluative judgement in the context of holistic competency development: Results of a Delphi study. *Assessment & Evaluation in Higher Education, 48*(4), 513-528. <https://doi.org/10.1080/02602938.2022.2088690>
- Luo, J., Chan, C. K. Y., & Zhao, Y. (2023). The development and validation of an instrument to measure evaluative judgement: A focus on engineering students' judgement of intercultural competence. *Assessment & Evaluation in Higher Education, 1-17*. <https://doi.org/10.1080/02602938.2023.2222937>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144. <https://doi.org/10.1007/BF00117714>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education, 76*, 467-481. <https://doi.org/10.1007/s10734-017-0220-3>

Note: The book was edited by David Boud, Rola Ajjawi, Phillip Dawson and Joanna Tai in 2018 and published by Routledge (ISBN: 978-1138089341).

Xiao Xie*

Universiti Putra Malaysia, Malaysia

Corresponding author: Xiao Xie

E-mail: gs58879@student.upm.edu.my

Received: 29/10/2023

Accepted: 14/01/2024

Published: 31/01/2024