

‘You Never Know If It’s Correct:’ An Exploration of Students’ Understandings on Evaluative Judgement in the Era of GenAI

Kubert Tianhang
WANG

Monica Xinjie
HUANG

Beatrice C.B.
CHU

Esme Dinelle
ANDERSON

Yonathan Aklilu
KIDANEMARIAM

Purpose of Study

Generative artificial intelligence (GenAI) is reshaping students’ learning in higher education through making it easy to generate seemingly high-quality outputs of students’ learning tasks and receive instant feedback (Bearman et al., 2024). This recent development raised a core educational concern about how students learn to make evaluative judgements about the quality of work and their epistemic stance (Tai et al., 2018; Walton et al., 2025). In higher education teaching and learning, evaluative judgement is central to ‘assessment-as-learning’ because students are expected to produce their work and also be able to judge quality in contextually appropriate and defensible ways (Yan & Boud, 2021).

Recent scholarly discussions have begun to address evaluative judgement in the era of GenAI, including how assessment practices might be adapted to GenAI-mediated teaching and learning environments (Corbin et al., 2025; Xia et al., 2024). However, there is still limited evidence on how students conceptualise ‘quality work’ and ‘defensible judgement strategies’ in the era of GenAI, and how they negotiate responsibility for judgement when GenAI plays different roles in higher education settings.



Theoretical Background

This study draws on a conceptualisation of evaluative judgement as a multidimensional construct comprising knowledge, attitude, capability, action, and identity (Luo & Chan, 2023). This makes the framework particularly suitable for contextualising GenAI-driven student learning, because the extension of the narrower focus of the original framework on evaluative judgement as merely a matter of ‘capability’ and offers a more comprehensive account of how learners understand the quality of work, value judgement, enact evaluative practices, and position themselves in relation to their own learning and broader assessment practices.



Research Design



This study adopts an exploratory qualitative approach using the critical incident technique (CIT) to understand experience-near accounts of students’ evaluative judgement in the GenAI learning environment (Merriam & Tisdell, 2016).



10 participants were recruited across disciplines reflecting the context-dependent nature of quality standards and evaluative practices in higher education (Bearman et al., 2024).



Interviews were conducted for 60-90 minutes for each student to reflect on their contextual mapping of disciplinary tasks and GenAI use, narration of three theoretically targeted incident types, and reflective prompts informed by our theoretical dimensions (Luo & Chan, 2023).

The three incident types serve as theory-probing contexts:

Incident A	A GenAI-assisted output initially appeared high quality but was later questioned, revised or refined.
Incident B	The student judged the quality of their own Gen-AI use process.
Incident C	AI feedback/evaluation was compared with the judgement from their own, their teachers and their peers, prompting refinement, acceptance or rejection of using the GenAI outputs.

Findings

Preliminary findings suggest that students’ evaluative judgement is enacted as reference system-based judgement in which learners assemble, prioritise and coordinate multiple references including disciplinary norms, teachers feedback, online resources and AI feedback to produce defensible evaluative strategies. This study refines Luo and Chan’s five-dimensional conceptualisation of evaluative judgement by providing more nuanced elaborations:

Knowledge

Knowledge extends beyond knowing criteria to include awareness of traceability and sensitivity to ‘AI Hallucinations’.

Action

Action becomes more process-driven comprising verification strategies alongside prompting governance and stopping decision when iterating the conversations with GenAI.

Identity

It is noteworthy that identity is restructured from a general assessor into a positioning of judgement authority and responsibility boundaries in relation to GenAI. Across incident types, these patterns point to an amplified identity-action coupling in the evaluative judgement under the GenAI era. Identity positioning appears to determine whether verification strategies are activated, thereby shaping whether judgement responsibility is retained or outsourced.

Originality/Value of Paper

This study contributes to the emerging scholarly discussions on the conceptualisation of evaluative judgement in the era of GenAI by providing students empirical understanding on evaluative judgement. It enriches current scholarly literatures on design-based recommendations by examining how students might reposition the relationship among the five dimensions of the evaluative judgement framework. This study also demonstrated the value of the CIT approach for theorising evaluative judgement in GenAI-related learning contexts without relying on researcher-imposed task materials.

References



Corresponding author: Kubert Tianhang WANG
Corresponding author’s email: kubertwang@ust.hk

Acknowledgements

This study is developed as part of the UGC project ‘Fund for Innovative Technology-in-Education: Fostering Digital Competency in the GenAI-GenZ Era: Utilizing a Student-Driven Community of Practice to Developing Teaching and Learning Practices.’ The authors gratefully acknowledge the support of this funding initiative and the contributions of participating students and colleagues involved in the project.