Impact of AI Support with Different Level of Learner Agency on Behavioral Engagement in Providing and Uptaking Peer Feedback

Xinmeng Hou[1], Wenli Chen[1*], Lishan Zheng[1], Guo Su[1], Xuanyu Chen[2], Qianru Lyu[3]

[1] National Institute of Education, Nanyang Technological University, Singapore

[2] Institute of Higher Education, Fudan University, China

[3] Human-Computer Interaction Institute, Carnegie Mellon University, USA

[*] wenli.chen@nie.edu.sg

Abstract: *This study examined how AI support with varying levels of learner agency influences learners' behavioral engagement during peer feedback construction and uptake. Ninety university students in 30 triads were randomly assigned to three conditions: Control (no AI support), AI-directed (AI which provides suggsted revision on learning artefacts), or AI-assisted (AI provides guidance for learners to improve learning artefacts). A conversation analysis framework was proposed for analyzing learners' behavioral engagement in human-AI dialogue through four dimensions: elaboration level, topic pursuit, difficulty response, and repair pursuit intensity. Analysis of the two AI conditions (n = 60) revealed task-dependent patterns: participants showed greater response depth and topic extension during feedback construction, but more clarification-seeking during feedback uptake. Significant interaction effects emerged between AI support type and learning task. AI-assisted learners displayed adaptive engagement, while AI-directed learners maintained consistent, undifferentiated engagement. These findings suggest that AI in an assistant role preserves learner autonomy and agency, and fosters task-adaptive engagement, while AI in a directive role may produce engagement patterns consistent with automation complacency.*

Keywords: AI-supported collaborative learning, peer feedback, behavioral engagement, conversation analysis, learner agency

1. Introduction

Feedback has long been considered a key factor influencing learning and achievement (Black & Wiliam, 1998; Hattie & Timperley, 2007). Effective feedback helps learners understand goals, compare their current performance to standards, and take action to close the gap (Sadler, 1989). Peer feedback has emerged as a scalable approach to delivering formative assessment, with reviews indicating adequate reliability and validity (Li et al., 2020; Topping, 1998). The concept of feedback literacy further emphasizes learners' active role in making sense of and acting upon feedback information (Carless & Boud, 2018). Beyond assessment accuracy, peer feedback may help develop learners' metacognitive skills and evaluative judgment (Nicol & Macfarlane-Dick, 2006). However, learners often struggle to provide constructive comments without adequate support (Yu & Lee, 2016), prompting exploration of scaffolding approaches (Min, 2005; Gielen & De Wever, 2015).

Generative AI has introduced new possibilities for feedback in education. Research distinguishes between AI-generated feedback and AI-supported peer feedback (Guo, 2024; Steiss et al., 2024), with hybrid human-AI models preserving cognitive benefits while addressing quality concerns (Banihashem et al., 2024; Kaliisa et al., 2025). However, providing and uptaking feedback involves fundamentally different learning mechanisms—feedback providers engage in evaluative judgment, while feedback receivers focus on interpretation and implementation (Nicol et al., 2014; Wu & Schunn, 2021)—yet few studies have examined how engagement patterns differ across these activities. Additionally, there is limited research conceptualizing AI's role by considering how much agency learners retain. AI can function as a directive agent or collaborative partner preserving learner agency (Darvishi et al., 2024), and this distinction

matters: human-centered AI fosters deeper learners' self-regulation and critical thinking, whereas directive approaches risk learners' passive dependency (Sperber et al., 2025; Zhan & Yan, 2025).

Despite growing interest, there is hardly any study examining learners' micro-level behavioral engagement across different feedback tasks—specifically, whether observable effort in dialogue differs between feedback construction and feedback uptake. In addition, no study has empirically compared how directive versus assistive AI support shapes moment-to-moment engagement patterns across these tasks. The present study addresses these gaps by: (1) proposing a novel conversation analysis (CA) framework for operationalizing learners' behavioral engagement in human-AI dialogue; and (2) providing empirical evidence of how the interaction between AI support type and peer feedback task type shapes learners' engagement patterns in human-AI dialogue.

2. Literature Review

Learning engagement is a multidimensional construct comprising behavioral, emotional, and cognitive components (Fredricks et al., 2004). Among these, behavioral engagement—defined as the effort, persistence, and attention students direct toward learning tasks (Skinner & Belmont, 1993; Finn & Zimmer, 2012)—has been identified as the strongest predictor of academic achievement (Finn & Zimmer, 2012). Linnenbrink and Pintrich (2003) further characterized behavioral engagement as encompassing effort, persistence, and instrumental help-seeking. Within feedback research specifically, Ellis (2010) conceptualized behavioral engagement as how learners respond to and act upon feedback, while Han and Hyland (2015) operationalized it through revision operations and observable strategies. This effort-centered view provides the foundation for examining how learners invest themselves in feedback-related tasks.

Critically, the demands placed on learners differ depending on which feedback activity they perform. Research on peer feedback has established that providing and receiving feedback involve fundamentally different learning mechanisms: feedback providers engage in evaluative judgment, problem detection, and criteria application, while feedback receivers focus on interpretation and implementation (Nicol et al., 2014; Patchan & Schunn, 2015). Wu and Schunn (2021) found that learning correlated more strongly with providing feedback than receiving it, suggesting that feedback construction activates deeper cognitive processing. This asymmetry likely stems from the evaluative demands of feedback provision, which require learners to articulate standards, identify discrepancies, and generate actionable suggestions (Lundstrom & Baker, 2009). Conversely, feedback uptake requires learners to interpret received comments, reconcile them with their own understanding, and implement revisions, which place different cognitive demands (Nelson & Schunn, 2009; Gielen et al., 2010; Chen et al., 2023). If these tasks impose distinct cognitive demands, they should also elicit qualitatively different patterns of behavioral engagement; yet this possibility has not been empirically examined.

The emergence of generative AI (Gen AI) has introduced new possibilities for scaffolding these feedback activities. Research distinguishes between AI-generated feedback and AI-supported peer feedback (Guo, 2024; Steiss et al., 2024), with hybrid human-AI models showing promise for preserving cognitive benefits while addressing quality concerns (Banihashem et al., 2024; Kaliisa et al., 2025). However, a key dimension remains underexplored: how much agency learners retain when interacting with AI. Darvishi et al. (2024) showed that AI can function as a directive agent delivering finished products or as a collaborative partner preserving learner agency. This distinction is consequential: human-centered AI that prioritizes learner agency may foster deeper self-regulation and sustained engagement, whereas directive approaches risk promoting passive dependency and automation complacency (Parasuraman & Manzey, 2010; Sperber et al., 2025; Zhan & Yan, 2025).

Despite these advances, two gaps persist. First, prior studies have typically treated feedback-related tasks as a single activity, without distinguishing between feedback construction and uptake—two activities that, as reviewed above, place distinct demands on learners and should therefore elicit different engagement patterns. Second, existing research conceptualizes AI involvement in binary terms (present or absent), without examining how different AI roles—

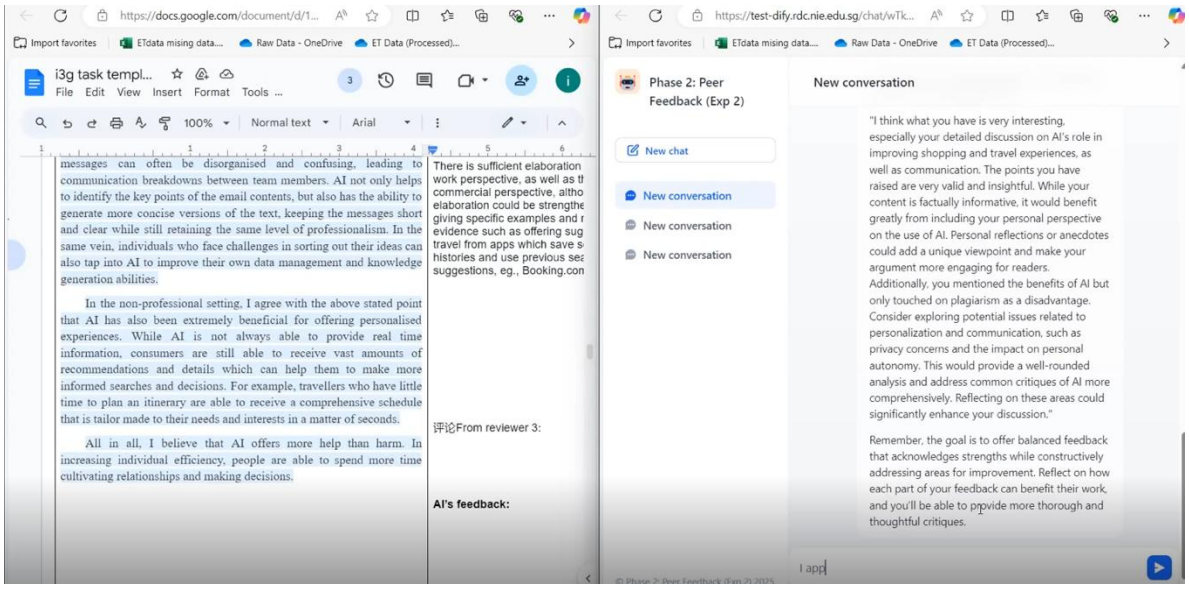
specifically, the degree of agency afforded to learners—shape the micro-level behavioral effort learners invest in human-AI dialogue across these distinct tasks. Building on this foundation, the present study addresses the following research questions: **RQ1.** How do learners' behavioral engagement patterns differ between feedback construction and feedback uptake when interacting with AI support? **RQ2.** Controlling for Peer feedback activity type, how does AI support type (directive vs. assisted) affect learners' behavioral engagement?

3. Methodology

Experiment design was adopted in this study to example the impact of AI support with different level of learner agency on learners' behavioral engagement in providing and uptaking peer feedback. The study employed a between-subjects randomized controlled design with three conditions—Control, AI-directed, and AI-assisted. Ninety students from a university in Singapore were organized into 30 triads (10 triads per condition). Participants completed a 30-minute protocol using a custom workspace integrating Google Docs with an embedded GPT-4o chatbot (see Figure 1). The control condition was included to establish a baseline for feedback quality; however, because control participants did not interact with the chatbot, the behavioral engagement analysis was restricted to the two AI supported peerr feedback conditions (n = 60). All claims should be interpreted as comparisons between AI support modalities, not AI-supported versus unsupported conditions.

Figure 1

AI-Supported Peer Feedback Workspace Interface



During feedback providing (15 minutes), AI-directed participants received AI-suggested reivisons on their learning artefacts to them to adopt or adapt; AI-assisted participants received guidance, and questions which scaffold them to iteratively refined the learning ratefacts. During feedback uptaking (15 minutes), participants revised essays based on received feedback. Table 1 summarizes prompt design differences.

Table 1

Comparison of AI prompt design across conditions and phases.

Phase	Condition	What AI Does	What Learner Does
Feedback	AI-Directed	Directly revise student's peer feedback for clarity and rubric alignment; marks changes in bold	Adapt AI-revised feedback for peers
	AI-Assisted	Offers suggestions, examples, and reflective questions to improve feedback	Implement all revisions independently
Feedforward	AI-Directed	Directly revises essay based on peer feedback; marks	Review and accept AI-

	changes in bold; no new evidence added	made revisions
AI-Assisted	Provides guidance on improving essay; includes examples and reflective questions	Make all revisions independently

To measure behavioral engagement, the study used conversation analysis (CA) methods for examining how participants demonstrate investment through observable discourse features (Stivers & Rossano, 2010; Schegloff, 2007). "Behavioral engagement" here refers to the effort, persistence, and active participation learners demonstrate in human-AI dialogue during learning tasks (Skinner & Belmont, 1993), operationalized through four CA-derived indicators, each mapping onto an established engagement dimension. Elaboration level captures effort and concentration: learners who produce extended turns with examples, reasoning, or synthesis demonstrate greater investment than those producing minimal acknowledgment tokens (Schegloff, 1982; Skinner & Belmont, 1993). Topic pursuit captures persistence: actively developing and extending conversational topics reflects sustained behavioral involvement rather than passive topic acceptance (Button & Casey, 1985; Fredricks et al., 2004; Connell & Wellborn, 1991). Difficulty response operationalizes orientation toward challenge, distinguishing learners who engage with difficult content from those who avoid it—a core distinction in engagement research (Schegloff et al., 1977; Skinner et al., 2009). Repair pursuit intensity captures instrumental help-seeking: the extent to which learners actively seek clarification and verification reflects purposeful engagement with understanding rather than passive acceptance of AI-provided content (Schegloff, 2007; Linnenbrink & Pintrich, 2003).

Collectively, higher levels across these indicators may also serve as behavioral proxies for learner agency—the capacity to intentionally influence one's own learning through proactive, self-regulated action (Reeve & Tseng, 2011). However, we note that our primary construct is behavioral engagement; inferences about agency remain secondary and should be interpreted with appropriate caution, as these indicators do not directly measure intentionality or self-regulation. This distinction is important given concerns that AI assistance may promote dependency rather than independent skill development (Darvishi et al., 2024). Each indicator was coded on a four-point ordinal scale (0–3), drawing on CA principles for identifying interactional phenomena (Sidnell & Stivers, 2013) while adapting them for systematic quantitative analysis. This approach follows precedents in applied linguistics and interaction research that have operationalized CA-derived constructs for larger-scale comparative studies (Stivers, 2015; Stivers & Rossi, 2025). The unit of analysis was the individual learner turn, defined as each discrete message a learner sent to the AI chatbot during the interaction. A conversation comprised the full dialogue sequence between one learner and the chatbot within a single learning task. This yielded 363 coded turns across 120 conversations (60 learners under AI-supported conditions × 2 tasks). Two trained coders independently coded all learner turns, with inter-rater reliability assessed using Cohen's kappa (see Table 2). The coding schema demonstrated strong agreement across indicators, with elaboration level achieving near-perfect reliability ($\kappa = .95$) and the remaining indicators showing substantial agreement (κ range = .80–.66), consistent with benchmarks for conversational data analysis (Landis & Koch, 1977).

Table 2

Inter-rater reliability for behavioral engagement indicators.

Indicator	Description	Scale	κ
Elaboration Level	Degree of substantive depth in learner contributions	0–3	.95
Topic Pursuit	Degree of active topic development and extension	0–3	.80
Difficulty Response	Degree of orientation toward task complexity	0–3	.77
Repair Pursuit Intensity	Degree of active verification and understanding seeking	0–3	.66

Note. κ = Cohen's kappa calculated between two coders. Adjacent agreement reflects the percentage of annotations within one scale point.

4. Findings

This study conducted nonparametric analyses using Mann-Whitney U tests for pairwise comparisons and Kruskal-Wallis H tests for interaction effects across 120 conversations comprising 363 turns. Effect sizes are reported using rank-biserial correlation (r), with .10, .30, and .50 as small, medium, and large effects (Cohen, 1988).

As shown in Table 3, three of four indicators showed significant task differences. Elaboration level and topic pursuit were both significantly higher during feedback construction, meaning students provided more detailed responses and more actively extended conversation topics when giving peer feedback. The task difference in elaboration level represented a small-to-medium effect (r = .20). Difficulty response showed no significant difference, possibly because the 30-minute protocol with a relatively accessible essay topic did not elicit sufficient variation in challenge orientation, or because both AI conditions provided enough scaffolding to compress the range of difficulty response behaviors. However, the pattern reversed for repair pursuit intensity, which was significantly higher during feedback uptake (medium effect, r = .35), indicating that students asked more clarifying questions and sought more verification when revising their own essays—the most practically meaningful task difference observed.

Table 3

Behavioral engagement indicators across experimental phases.

Indicator	Feedback providing	Feedback Uptaking	Difference	*U*	*p*	r
Elaboration Level	2.02	1.62	0.40	19467.00	.010**	.20
Topic Pursuit	2.14	1.96	0.18	18984.00	.008**	.18
Difficulty Response	1.36	1.22	0.13	17422.31	.100	.11
Repair Pursuit Intensity	0.31	0.64	−0.33	13465.00	<.001***	.35

Note. Diff = Phase 2 − Phase 3. Effect size r = rank-biserial correlation. *p < .01. ***p < .001.

Kruskal-Wallis H tests revealed significant interaction effects for three of four indicators (Table 4). For elaboration level and topic pursuit, AI-assisted students showed substantial declines from construction to uptake, while AI-directed students remained stable. For repair pursuit intensity, AI-assisted students dramatically increased clarification-seeking during uptake—nearly a fourfold increase—while AI-directed students showed minimal change. Difficulty response showed no significant interaction.

Table 4

Interaction effects between phase and AI support condition.

Indicator	providing + Assisted	providing + Directed	Uptake + Assisted	Uptake + Directed	*H*	*p*
Elaboration Level	2.35	1.74	1.51	1.73	25.82	<.001***
Topic Pursuit	2.37	1.94	1.89	2.03	18.47	<.001***
Difficulty Response	1.51	1.23	1.21	1.24	5.37	.147
Repair Pursuit Intensity	0.21	0.40	0.77	0.50	21.94	<.001***

Note. H = Kruskal-Wallis test statistic with df = 3. ***p < .001. Construction = feedback construction phase; Uptake = feedback uptake phase; Assisted = AI-assisted condition; Directed = AI-directed condition

These interaction effects reveal that the two AI support conditions shaped behavioral engagement differently across tasks. In the AI-assisted condition, the scaffolded approach—providing guidance and reflective questions rather than finished products, elicited task-dependent behaviors: learners produced more elaborate contributions during peer feedback but shifted toward clarification-seeking during essay revision, actively verifying their understanding when responsible for implementing changes themselves. In contrast, the AI-directed condition showed stable engagement

across tasks, likely because students had less need to elaborate or seek clarification when the AI had already completed the substantive revision work. This pattern suggests that the degree of learner agency preserved by the AI support design fundamentally shapes how learners distribute their effort across different feedback activities.

5. Discussion and Conclusion

By adapting conversation analysis to operationalize engagement in human-AI dialogue, this study investigated how different AI support levels influence behavioral engagement during peer feedback activities.

Task-dependent engagement. The differential patterns suggest feedback construction and uptake elicit qualitatively different learner investment. During construction, higher elaboration and topic development reflect the evaluative demands of providing feedback (Nicol et al., 2014; Lundstrom & Baker, 2009). During feedback uptake, higher clarification-seeking reflects interpretive demands of reconciling received feedback with one's own understanding (Nelson & Schunn, 2009). These findings provide micro-level empirical support for the theoretical distinction between feedback provision and uptake demands (Wu & Schunn, 2021).

AI role and adaptive engagement. The significant interaction effects reveal that the two AI conditions shaped engagement in qualitatively different ways, a pattern interpretable through self-determination theory (SDT; Ryan & Deci, 2000). The AI-assisted condition—which provided guidance and reflective questions rather than finished products—supported learner autonomy by preserving space for independent decision-making. This autonomy support manifested as adaptive engagement: learners invested in elaboration during feedback construction (when the task demanded evaluative generation) but shifted toward verification during feedback uptake (when the task demanded interpretive implementation). Such task-responsive flexibility suggests that scaffolded AI support enabled learners to calibrate their effort to the specific demands of each activity, consistent with SDT's prediction that autonomy support fosters more self-regulated and intrinsically motivated behavior. In contrast, the AI-directed condition produced stable, undifferentiated engagement across tasks, aligning with the automation complacency framework (Parasuraman & Manzey, 2010): when AI provides finished products that require only review and acceptance, learners may reduce their active cognitive processing regardless of task demands. The consistency of engagement in the AI-directed condition (neither increasing during demanding tasks nor decreasing during simpler ones) suggests that the directive AI may have attenuated the natural variation in effort that different feedback activities should elicit. While consistent engagement might appear desirable on the surface, it may reflect a ceiling effect on passivity rather than genuine stability of effort.

Design implications. AI should be positioned as a dialogic partner that scaffolds learners' thinking through questions and suggestions rather than delivering ready-made solutions. Support should be task-sensitive: during feedback construction, AI should prompt evaluative reasoning and criteria application; during feedback uptake, AI should support interpretation and encourage learners to verify their understanding before implementing changes. Third, designers should monitor for engagement uniformity as a potential warning sign of automation complacency, building in mechanisms that require active learner processing, such as requiring learners to articulate their reasoning before receiving AI suggestions.

Limitations. Several limitations warrant consideration. First, the 30-minute protocol captures behavioral engagement during an initial exposure to AI-supported feedback. Because engagement patterns can evolve as novelty effects diminish and learners develop stable interaction strategies with AI tools, the adaptive engagement patterns observed in the AI-assisted condition may not persist over longer interventions. Longitudinal designs are needed to examine the stability of these patterns. Second, our behavioral engagement measures capture only one dimension of the multidimensional engagement construct; future research should integrate emotional and cognitive engagement measures to provide a more complete picture. Third, the behavioral analysis was restricted to the two AI conditions because the control group did not interact with the chatbot; future studies should incorporate cross-condition measures (e.g., feedback quality, writing improvement) to examine whether the engagement differences observed here translate

into differential learning outcomes. Finally, the single-university sample from Singapore limits generalizability across educational and cultural contexts.

Acknowledgements

This study was funded by the NIE Research Support for Senior Academic Administrators (RS-SAA) Grant (#022028-00001) supported by the National Institute of Education & Nanyang Technological University, Singapore. This study was approved by the NTU IRB (IRB-2022-210).

References

Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? International Journal of Educational Technology in Higher Education, 21(1), Article 23.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education, 5(1), 7–74.

Button, G., & Casey, N. (1985). Topic nomination and topic pursuit. Human Studies, 8(1), 3–55.

Carless, D., & Boud, D. (2018). The development of student feedback literacy. Assessment & Evaluation in Higher Education, 43(8), 1315–1325.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness. In M. R. Gunnar & L. A. Sroufe (Eds.), Minnesota Symposia on Child Psychology (Vol. 23, pp. 43–77). Lawrence Erlbaum.

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. Computers & Education, 210, Article 104967.

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. Studies in Second Language Acquisition, 32(2), 335–349.

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson et al. (Eds.), Handbook of research on student engagement (pp. 97–131). Springer.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59–109.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. Learning and Instruction, 20(4), 304–315.

Gielen, M., & De Wever, B. (2015). Structuring the peer assessment process. Journal of Computer Assisted Learning, 31(5), 435–449.

Guo, K. (2024). EvaluMate: Using AI to support students' feedback provision in peer assessment for writing. Assessing Writing, 61, Article 100864.

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback. Journal of Second Language Writing, 30, 31–44.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112.

He, W., & Gao, Y. (2023). Explicating peer feedback quality and its impact on feedback implementation in EFL writing. Frontiers in Psychology, 14, Article 1177094.

Kaliisa, R., Misiejuk, K., López-Pernas, S., & Saqr, M. (2025). How does AI compare to human feedback? Educational Psychology. Advance online publication.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2020). Does peer assessment promote student learning? Assessment & Evaluation in Higher Education, 45(2), 193–211.

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement. Reading & Writing Quarterly, 19(2), 119–137.

Lundstrom, K., & Baker, W. (2009). To give is better than to receive. Journal of Second Language Writing, 18(1), 30–43.

Min, H.-T. (2005). Training students to become successful peer reviewers. System, 33(2), 293–308.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback. Instructional Science, 37(4), 375–401.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning. Studies in Higher Education, 31(2), 199–218.

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education. Assessment & Evaluation in Higher Education, 39(1), 102–122.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation. Human Factors, 52(3), 381–410.

Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback. Instructional Science, 43(5), 591–614.

Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement. Contemporary Educational Psychology, 36(4), 257–267.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation. American Psychologist, 55(1), 68–78.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144.

Schegloff, E. A. (1982). Discourse as an interactional achievement. In D. Tannen (Ed.), Analyzing discourse (pp. 71–93). Georgetown University Press.

Schegloff, E. A. (2007). Sequence organization in interaction. Cambridge University Press.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in repair. Language, 53(2), 361–382.

Sidnell, J., & Stivers, T. (Eds.). (2013). The handbook of conversation analysis. Wiley-Blackwell.

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom. Journal of Educational Psychology, 85(4), 571–581.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection. Educational and Psychological Measurement, 69(3), 493–525.

Sperber, L., MacArthur, M., Minnillo, S., Stillman, N., & Whithaus, C. (2025). Peer and AI Review + Reflection (PAIRR). Computers and Composition, 76, Article 102921.

Steiss, J., et al. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. Learning and Instruction, 91, Article 101894.

Stivers, T., & Rossano, F. (2010). Mobilizing response. Research on Language and Social Interaction, 43(1), 3–31.

Stivers, T. (2015). Coding social interaction. Research on Language and Social Interaction, 48(1), 1–19.

Stivers, T., & Rossi, G. (2025). Finding codability. Research on Language and Social Interaction, 58(3), 240–257.

Topping, K. (1998). Peer assessment between students in colleges and universities. Review of Educational Research, 68(3), 249–276.

Wu, Y., & Schunn, C. D. (2021). The effects of providing and receiving peer feedback. American Educational Research Journal, 58(3), 492–526.

Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). Language Teaching, 49(4), 461–493.

Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback. Assessment & Evaluation in Higher Education. Advance online publication.