EDUVELOP
JOURNAL OF ENGLISH EDUCATION AND DEVELOPMENT
Nationally Accredited Journal Decree No.10/C/C3/DT.05.00/E5/2025 (Sinta 2)

**Volume 9, No. 1,** March 2026 • **Page: 286 - 296**
**DOI:** https://doi.org/v9i1.6166
**ISSN** 2597-713X (print)
**ISSN** 2597-7148 (online)

**OPEN ACCESS: Research Article**

# Formative Assessment Practiced in Higher Education: Lecturers' and Students' Perceptions

Achmad Taqlidul Chair Fachruddin[1*]

[1*]English Education Department, Sekolah Tinggi Agama Islam Negeri Majene, Majene, Indonesia

*Correspondence e-mail: achmadtaqlidulchair@stainmajene.ac.id

| | |
|---|---|
| **Received :** 2026-01-07 | **Accepted :** 2026-02-10 |
| **Revision :** 2026-03-10 | **Published :** 2026-03-31 |

**Abstract**
Discrepancies between lecturers' and students' perceptions of formative assessment can weaken instructional alignment and reduce learning effectiveness in higher education. This study aimed to (1) describe formative assessment practices at STAIN Majene and (2) compare lecturers' and students' perceptions across five dimensions: learning objectives, monitoring, feedback, self-assessment, and peer assessment. Using a descriptive quantitative, cross-sectional design, questionnaire data were collected from 20 lecturers and 250 odd-semester students at STAIN Majene. The instrument consisted of 25 items and demonstrated acceptable internal consistency across dimensions, with combined Cronbach's alpha coefficients ranging from .70 to .83. Descriptive findings revealed that lecturers rated learning objectives ($M = 4.38$, $SD = 0.533$) and monitoring ($M = 4.02$, $SD = 0.536$) more positively, while students rated feedback most highly ($M = 4.40$, $SD = 0.554$). Inferential analysis using Welch's independent-samples t test showed significant differences in learning objectives, $t(21.71) = 8.115$, $p < .001$, and feedback, $t(24.51) = -10.289$, $p < .001$, while monitoring and peer assessment did not differ significantly. Self-assessment showed a smaller difference at the .05 level, $t(24.51) = -2.374$, $p = .026$. Overall, the findings suggest that formative assessment at STAIN Majene remains stronger in lecturer-led practices than in learner-activating dimensions, particularly peer assessment. As an institutional study situated in a state Islamic higher education context in West Sulawesi, this research contributes empirical evidence on lecturer–student perceptual gaps and offers directions for strengthening learning objective transparency, sustaining effective feedback, and systematizing self- and peer-assessment practices.

**Keywords:** *Formative Assessment; Lecturer–Student Perceptions; Learning Objectives; Feedback; Self-Assessment; Peer Assessment; STAIN Majene*

DOI: https://doi.org/v9i1.6166

## 1. Introduction

Formative assessment (FA) refers to evidence-gathering processes embedded in instruction that are used to adjust teaching and support students' next learning steps (Sadler, 1989; Black & Wiliam, 1998; Black & Wiliam, 2004). Oswalt (2013) stated that formative assessment consists of learning targets, monitoring, feedback, peer assessment, and self-assessments. In higher education, FA is expected to strengthen student-centred learning by making learning goals explicit, monitoring progress toward those goals, and providing feedback that students can use to regulate their learning. In this study, FA is operationalised through five interconnected dimensions: (1) learning objectives (sharing and using targets/criteria), (2) monitoring (eliciting evidence of learning), (3) feedback (information that guides improvement), (4) self-assessment (structured learner reflection against criteria), and (5) peer assessment (peer review and peer feedback) (Oswalt, 2013). However, the effectiveness of FA depends not only on what lecturers intend to do but also on what students actually perceive and use. Prior studies show that lecturers and students may interpret the same assessment practices differently, and these perception gaps can reduce feedback uptake, weaken learner agency, and undermine assessment transparency (Hansen, 2020; van der Kleij, 2019; Veugen et al., 2021). In higher education, this issue is closely linked to students' feedback literacy and the degree to which assessment criteria, task purposes, and feedback expectations are made visible and actionable in everyday teaching (Evans, 2013; Nicol, 2010; Jonsson, 2013). The elements of this formative assessment are used continually for teaching and learning improvement known as an instructional correction (TEAL, 2022). The continuity of the instructional corrective identified in higher education will lead to student-centered learning and lifelong learning. Hoidn (2016) advocacy for learner-centered practices reflects a significant pedagogical shift from traditional teacher-fronted instruction to a more constructivist model that empowers students. This approach rightly positions learners as active agents in their own intellectual development, emphasizing that the cultivation of metacognitive skills such as critical thinking, autonomy, and reflection is as crucial as the acquisition of content knowledge itself.

There have been few studies on students' and lecturers' perceptions of how formative assessment practices in the classroom (Kyaruzi et al., 2019; Mahmood & Ghaleb, 2024; Hill & Edwards, 2019). Specifically, Veugen et al. (2021) found that students and teachers perceived formative assessment differently depending on their backgrounds. In his research, Van der Kleij (2019) found that students with higher achievers have more positive perceptions of their teachers' feedback than students with lower achievers. Christoforidou & Kyriakides (2021) also found that the more teaching experience, the more formative assessment was used. Despite increasing emphasis on FA in higher education, evidence from Indonesian higher education—particularly in smaller or developing institutions—is still limited. STAIN Majene represents a relevant local context to examine whether FA practices are experienced as intended by lecturers and whether students interpret those practices as meaningful learning support.

DOI: https://doi.org/v9i1.6166

Accordingly, this study has two objectives: (1) to describe lecturers' and students' perceptions of formative assessment practices across the five dimensions, and (2) to identify areas of alignment and discrepancy between both groups' perceptions. By mapping perception profiles, the study provides empirical direction for improving assessment transparency (learning objectives and monitoring), sustaining effective feedback practices, and designing feasible self- and peer-assessment routines in higher education classrooms.

## 2. Methods

This study used a descriptive quantitative design to capture and compare lecturers' and students' perceptions of formative assessment practices. A survey approach was appropriate because the target construct is perceptual (self-reported experience) and the study aims to identify patterns of alignment/discrepancy between two respondent groups. Meanwhile, a descriptive approach was utilized to systematically and accurately describe the current state of perceptions regarding formative assessment practices at STAIN Majene, without any manipulation of variables.

The target population comprised all lecturers (N = 189) and students (N = 2,426) at STAIN Majene. To improve representativeness and reduce selection bias, the study applied a stratified sampling logic. For lecturers, strata were defined by faculty/study program, while for students, strata were defined by study program and semester grouping (odd semesters). Within each stratum, participants were invited proportionally to the population distribution, and recruitment continued until the minimum sample size per stratum was reached. This approach was adopted to increase the generalizability of findings within the STAIN Majene context. A total of 20 lecturers and 250 odd-semester students participated in the study. Because the lecturer sample is relatively small compared to the population.

Data were collected using a closed-ended questionnaire with a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). The instrument consisted of 25 items adapted from Oswalt's (2013) formative assessment components and organised into five dimensions: learning objectives, monitoring, feedback, self-assessment, and peer assessment. Items were written to capture routine classroom practices (e.g., clarity of objectives, use of monitoring evidence, usefulness of feedback, opportunities for reflection, and peer review). Because the instrument was researcher-administered and context-specific, the study reports internal consistency reliability and treats broader validity evidence as a limitation to be addressed in future research. The questionnaire was administered online via Google Forms. Participation was voluntary; respondents were informed about the study purpose, anonymity, and the use of aggregated results for research reporting.

Responses were exported from Google Forms, screened for completeness, and analysed using SPSS (Version 25). Internal consistency reliability was assessed using Cronbach's Alpha for each dimension (Table 1). For each dimension, item scores were aggregated and mean scores and standard deviations were computed for lecturers and students. To represent alignment/discrepancy, mean gaps (lecturer mean minus

DOI: https://doi.org/v9i1.6166

student mean) were calculated and interpreted alongside descriptive results.

With the instrument's reliability established, the primary analysis was conducted in several stages. Initially, a frequency analysis was performed to calculate the frequencies and percentages for each of the 25 individual statement items. This step was crucial for understanding the basic distribution of responses across the Likert scale for both the lecturer and student groups. Subsequently, the item scores within each of the five dimensions (sharing learning targets, monitoring, feedback, self-assessment, and peer assessment) were aggregated to compute a composite mean score for each dimension. This mean score represents the central tendency of perception for that specific construct. To further enrich the analysis, the standard deviation was calculated alongside each mean. This provided a vital measure of variability, indicating how tightly clustered or widely spread the perceptions were around the average score. In addition to descriptive statistics, inferential analysis was carried out to determine whether the perceptual differences between lecturers and students were statistically significant across the five dimensions of formative assessment. Given the unequal group sizes, Welch's independent-samples t-test was employed. To complement significance testing, effect sizes were computed using Hedges' g in order to estimate the magnitude of the differences. A significance level of .05 was applied in all comparisons. All resulting statistical outputs were then organized into tables to facilitate a systematic and clear comparison of the perception profiles between lecturers and students, directly addressing the core aims of this research.

## 3. Results and Discussions

### 3.1. Result

This section reports descriptive results of the five formative assessment dimensions. Reliability estimates for each dimension were acceptable for exploratory descriptive research, ranging from .64 to .83 across groups (Table 1). The relatively lower alpha for peer assessment among students is consistent with its low mean score and may indicate that peer assessment is not implemented consistently enough for students to form stable judgements about it.

**Table 1.** Cronbach's Alpha scores for each aspect

| Formative Assessment Aspects | α - Students | α - Lecturers | α - Students and Lecturers | Interpretation |
|---|---|---|---|---|
| Learning Objectives | .78 | .73 | .76 | Lecturers see strong clarity & relevance; students may not fully grasp or apply objectives. |
| Monitoring | .75 | .63 | .83 | Perceptions close, but students may see less individualized monitoring. |
| Feedback | .79 | .69 | .81 | Students highly value feedback; lecturers underestimate its impact. |
| Self-assessment | .77 | .77 | .82 | Students find some benefit; lecturers feel it's not well-structured. |

DOI: https://doi.org/v9i1.6166

| | | | | |
|---|---|---|---|---|
| Peer-assessment | .64 | .69 | .70 | Both rate low; indicates underutilization and possible training needs. |

This study examined lecturers' and students' perceptions of formative assessment practices across five dimensions: learning objectives, monitoring, feedback, self-assessment, and peer assessment. Descriptively, the findings show both convergence and discrepancy between the two groups. Among lecturers, the highest mean score was found for learning objectives (M = 4.38, SD = 0.533), followed by monitoring (M = 4.02, SD = 0.536), indicating that lecturers generally perceived these practices as being implemented strongly in classroom instruction. By contrast, students reported their highest mean score in feedback (M = 4.40, SD = 0.554), followed by monitoring (M = 3.88, SD = 0.513), suggesting that students viewed feedback as the most salient and beneficial aspect of formative assessment in their learning experience. Self-assessment received moderate ratings from both groups (lecturers: M = 2.96, SD = 0.424; students: M = 3.20, SD = 0.554), whereas peer assessment received the lowest ratings overall (lecturers: M = 2.52, SD = 0.495; students: M = 2.40, SD = 0.468), indicating that this dimension remained the least developed in classroom practice.

To extend the descriptive findings, inferential analysis was conducted using Welch's independent-samples t test because of the unequal group sizes between lecturers and students. The results revealed a statistically significant difference in learning objectives, $t(21.71) = 8.115$, $p < .001$, $g = 2.003$, with lecturers reporting substantially higher perceptions than students. A significant difference was also found in feedback, $t(24.51) = -10.289$, $p < .001$, $g = -1.900$, although in this case students rated feedback significantly higher than lecturers. In self-assessment, the difference between groups was smaller but still statistically significant at the .05 level, $t(24.51) = -2.374$, $p = .026$, $g = -0.438$, with students again reporting slightly more positive perceptions. By contrast, no statistically significant differences were found in monitoring, $t(21.88) = 1.128$, $p = .272$, $g = 0.271$, or peer assessment, $t(21.81) = 1.047$, $p = .306$, $g = 0.255$.

Taken together, these findings indicate that the most pronounced lecturer–student perceptual gaps were concentrated in learning objectives and feedback. Monitoring appeared relatively aligned between the two groups, while self-assessment and peer assessment remained weaker dimensions overall. Although self-assessment reached statistical significance at the conventional .05 level, the effect size was notably smaller than those found for learning objectives and feedback. Overall, the inferential results suggest that the main discrepancies lie not in formative assessment as a whole, but in how particular dimensions are enacted and experienced by lecturers and students.

**Table 2.** Comparison of Lecturers' and Students' Perceptions Across the Five Dimensions of Formative Assessment

| Formative Assessment Aspects | Lecturers M (SD) | Students M (SD) | *t* | *df* | *p* | Hedges' *g* | Interpretation |
|---|---|---|---|---|---|---|---|
| Learning Objectives | 4.38 (0.533) | 3.38 (0.495) | 8.115 | 21.71 | < .001 | 2.003 | Significant |

DOI: https://doi.org/v9i1.6166

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Monitoring | 4.02 (0.536) | 3.88 (0.513) | 1.128 | 21.88 | .272 | 0.271 | Not significant |
| Feedback | 3.36 (0.424) | 4.40 (0.554) | -10.289 | 24.51 | < .001 | -1.900 | Significant |
| Self-assessment | 2.96 (0.424) | 3.20 (0.554) | -2.374 | 24.51 | .026 | -0.438 | Significant |
| Peer-assessment | 2.52 (0.495) | 2.40 (0.554) | 1.047 | 21.81 | .306 | 0.255 | Not significant |

*Note:* Welch's independent-samples t test was used because of unequal group sizes. Positive effect sizes indicate higher lecturer ratings, whereas negative effect sizes indicate higher student ratings. Under a Bonferroni-adjusted threshold ($p < .01$), only learning objectives and feedback remained statistically significant.

## 3.2. Discussions

The findings demonstrate that formative assessment at STAIN Majene is not perceived uniformly across its different dimensions. Rather, the pattern of results suggests that some elements of formative assessment are more established and more visible than others. Interpreted through the framework of formative assessment, these results reinforce the view that effective formative practice involves more than the delivery of assessment activities; it also depends on whether students can recognize, interpret, and use those practices to regulate their own learning (Black & Wiliam, 2004; van der Kleij, 2019). In the present study, the lecturer-led dimensions of learning objectives and monitoring tended to be rated more positively by lecturers, whereas students responded most positively to feedback. At the same time, the learner-activating dimensions of self-assessment and peer assessment remained relatively weak, suggesting that formative assessment at STAIN Majene may still be more teacher-driven than fully participatory.

The most pronounced difference emerged in learning objectives. Lecturers rated this dimension very highly, whereas students reported a markedly lower mean, and the inferential analysis confirmed that this was the strongest gap between the two groups. This finding suggests that lecturers may believe they are communicating course goals clearly and aligning them effectively with instruction and assessment, but students do not always experience those objectives as sufficiently clear, accessible, or useful for guiding their learning. From a formative assessment perspective, this distinction is important because learning objectives are not merely statements of intention; they are meant to function as criteria that help learners understand what quality performance looks like and how they should direct their efforts (Black & Wiliam, 2004). In the context of STAIN Majene, this gap may reflect instructional routines in which objectives are formally presented but not consistently translated into student-friendly success criteria, exemplars, or rubrics. As a result, lecturers may perceive the objectives as already explicit, while students may still struggle to use them as practical learning guides.

DOI: https://doi.org/v9i1.6166

Feedback produced the opposite pattern. Students rated feedback significantly higher than lecturers, and the effect size was also very large. This is one of the most encouraging findings of the study because it indicates that students experience feedback as meaningful, timely, and useful for improving their performance. At the same time, lecturers' lower self-ratings may reflect a more self-critical stance toward their own practice, perhaps because they evaluate their feedback against ideal professional standards rather than against students' actual experience of it. This interpretation is consistent with previous scholarship emphasizing that feedback is most effective when it helps learners close the gap between current and desired performance (Van der Kleij, 2019). In this sense, the results suggest that feedback is already functioning as a relative strength within the STAIN Majene context. Nevertheless, the discrepancy also points to room for refinement, particularly by making feedback more explicitly dialogic and forward-looking, so that students not only receive comments but are also guided in how to act on them in subsequent learning tasks.

A smaller but statistically significant difference was also found in self-assessment, with students reporting slightly more positive perceptions than lecturers. Although this effect was much more modest than those found in learning objectives and feedback, it still suggests an interesting dynamic. Students may recognize some benefit in opportunities to reflect on their own strengths and weaknesses, even when lecturers themselves do not regard such opportunities as systematic or sufficiently structured. At the same time, because this difference is relatively small, it should be interpreted cautiously. What remains more important is the overall pattern: self-assessment was only moderate for both groups, indicating that it is not yet fully embedded in routine classroom practice. In formative assessment theory, self-assessment plays a central role in developing learners' capacity for self-regulation, judgment, and responsibility for learning (Black & Wiliam, 2004). The moderate scores in this study therefore suggest the need for stronger scaffolding, such as reflection prompts, checklists, and rubric-based self-review.

By contrast, monitoring and peer assessment did not differ significantly between lecturers and students. In the case of monitoring, both groups rated the dimension relatively positively, which implies a shared recognition that student progress is being tracked through quizzes, assignments, discussions, or similar classroom activities. The absence of a significant difference here is important because it suggests a degree of mutual perceptional alignment. However, alignment should not be confused with optimal implementation. Monitoring becomes genuinely formative only when the information collected is used responsively to shape instruction and provide tailored support. It is therefore possible that both lecturers and students recognize the presence of monitoring activities, while the pedagogical consequences of those activities remain less visible.

Peer assessment presents a different picture. Although there was no significant difference between lecturers and students, the means were low for both groups, making it the weakest dimension in the study. This indicates not a lack of agreement, but a shared perception that peer assessment is underutilized. Such a finding is consistent

DOI: https://doi.org/v9i1.6166

with the broader literature suggesting that peer assessment is often the most difficult dimension of formative assessment to institutionalize, despite its potential to foster evaluative judgment, collaboration, and deeper engagement with criteria (Black & Wiliam, 2004; Brookhart & McMillan, 2019). In the STAIN Majene context, the weak standing of peer assessment may be shaped by a combination of factors, including limited student familiarity with peer feedback, uncertainty about fairness, and classroom norms that make overt critique among peers less comfortable. Without clear structures, modeling, and low-stakes opportunities for practice, peer assessment may remain peripheral rather than becoming a meaningful part of classroom learning.

Overall, the inferential findings sharpen the interpretation of the descriptive results by showing that not every difference in mean scores reflects a statistically meaningful divergence. The most stable and substantial discrepancies appear in learning objectives and feedback, while monitoring is relatively aligned and peer assessment is jointly perceived as underdeveloped. This distinction matters because it helps identify where improvement efforts should be concentrated. Rather than treating all five dimensions as equally problematic, the findings suggest that STAIN Majene would benefit most from strengthening the transparency and usability of learning objectives, while simultaneously preserving and extending the positive feedback practices already recognized by students.

Several limitations need to be acknowledged. First, the lecturer sample was relatively small in relation to the overall lecturer population, and the sampling approach limits the broader generalizability of the findings. Second, the study relied on self-report questionnaire data, which capture perceptions rather than directly observed classroom behavior. As a result, the findings should be interpreted as an institutional snapshot of perceived formative assessment practices rather than as a definitive representation of all actual classroom processes. Even so, the study makes a useful contribution by demonstrating where lecturer and student perceptions converge, where they diverge, and which aspects of formative assessment appear to require the most immediate pedagogical attention in the local higher education context of STAIN Majene.

## 4. Conclusion

This study concludes that formative assessment at STAIN Majene is valued by both lecturers and students, but its implementation remains uneven across dimensions. Inferential findings show that the clearest lecturer–student differences occurred in learning objectives and feedback: lecturers rated learning objectives significantly higher, whereas students rated feedback significantly higher. Self-assessment showed a smaller difference, while monitoring and peer assessment did not differ significantly. Although both groups recognized monitoring positively, peer assessment remained the weakest dimension overall, indicating limited classroom use. These results suggest that formative assessment in this context is still stronger in lecturer-led practices than in learner-activating ones. Despite limitations related to sample imbalance and self-report data, the study provides useful institutional evidence and highlights the need to

strengthen objective transparency, maintain effective feedback, and develop more systematic self- and peer-assessment practices.

**Acknowledgement**

This research was supported by *Lembaga Penelitian dan Pengabdian kepada Masyarakat (LITABDIMAS)*, whose financial assistance is gratefully acknowledged. The author also extends appreciation to STAIN Majene for institutional support and to all participating lecturers and students for their invaluable contributions to the completion of this study.

**Conflicts of Interest**

The author declares no conflict of interest related to this research.

**References**

Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. Andrade & G. Cizek (Eds.), Handbook of formative assessment (pp. 90–105). Routledge.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. Teachers College Record, 106(14), 20–50. https://doi.org/10.1177/016146810410601402

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. Assessment & Evaluation in Higher Education, 31(4), 399–413. https://doi.org/10.1080/02602930600679050

Brookhart, S. M. (2008). How to give effective feedback to your students. ASCD.

Brookhart, S. M. (2013). How to create and use rubrics for formative assessment and grading. ASCD.

Brookhart, S. M., & McMillan, J. H. (2019). Classroom assessment and educational measurement. In Classroom assessment and educational measurement. Routledge. https://doi.org/10.4324/9780429507533

Carless, D. (2015). Excellence in university assessment: Learning from award-winning practice. Routledge.

Christoforidou, M., & Kyriakides, L. (2021). Developing teacher assessment skills: The impact of the dynamic approach to teacher professional development. Studies in Educational Evaluation, 70, 101051. https://doi.org/10.1016/j.stueduc.2021.101051

Evans, C. (2013). Making sense of assessment feedback in higher education. Review of Educational Research, 83(1), 70–120. https://doi.org/10.3102/0034654312474350

Hansen, G. (2020). Formative assessment as a collaborative act: Teachers' intention and students' experience. Studies in Educational Evaluation, 66, 100904. https://doi.org/10.1016/j.stueduc.2020.100904

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. https://doi.org/10.3102/003465430298487

Hill, L., & Edwards, F. (2019). Student perceptions of their involvement in formative assessment feedback practices: "I can do it myself." Assessment Matters, 13, 6–43. https://doi.org/10.18296/am.0036

Hoidn, S. (2016). The pedagogical concept of student-centred learning in the context of European higher education reforms. European Scientific Journal, 12(28), 439. https://doi.org/10.19044/esj.2016.v12n28p439

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. Active Learning in Higher Education, 14(1), 63–76. https://doi.org/10.1177/1469787412467124

Kyaruzi, F., Strijbos, J. W., Ufer, S., & Brown, G. T. L. (2019). Students' formative assessment perceptions, feedback use and mathematics performance in secondary schools in Tanzania. Assessment in Education: Principles, Policy & Practice, 26(3), 278–302. https://doi.org/10.1080/0969594X.2019.1593103

Mahmood, S. K., & Ghaleb, N. (2024). Kurdish EFL students' perceptions towards summative and formative assessment at Salahaddin University. Journal of Tikrit University for Humanities, 31(5), 38–62. https://doi.org/10.25130/jtuh.31.5.2024.25

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. Assessment & Evaluation in Higher Education, 35(5), 501–517. https://doi.org/10.1080/02602931003786559

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218. https://doi.org/10.1080/03075070600572090

Oswalt, S. G. (2013). Identifying formative assessment in classroom instruction: Creating an instrument to observe use of formative assessment in practice (Doctoral dissertation).

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions

for research. Frontiers in Psychology, 8, 422. https://doi.org/10.3389/fpsyg.2017.00422

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18, 119–144. https://doi.org/10.1007/BF00117714

Shute, V. J. (2008). Focus on formative feedback. Review of Educational Research, 78(1), 153–189. https://doi.org/10.3102/0034654307313795

Topping, K. J. (1998). Peer assessment between students in colleges and universities. Review of Educational Research, 68(3), 249–276. https://doi.org/10.3102/00346543068003249

van der Kleij, F. M. (2019). Comparison of teacher and student perceptions of formative assessment feedback practices and association with individual student characteristics. Teaching and Teacher Education, 85, 175–189. https://doi.org/10.1016/j.tate.2019.06.010

Veugen, M. J., Gulikers, J. T. M., & den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. Studies in Educational Evaluation, 70, 101027. https://doi.org/10.1016/j.stueduc.2021.101027

Wiliam, D. (2011). Embedded formative assessment. Solution Tree Press.