



Unpacking Student Responses to Discrepant Peer and Teacher Feedback: A Cross-National Comparison[☆]

Carolina Lopera-Oquendo ^{a,1,*}, Anastasiya A. Lipnevich ^{b,2}, Ligia Tomazin ^{a,3}, Ignacio Máñez ^{c,4}, Samuel P. León ^{d,5}, Nicola Beatson ^{e,6}

^a The Graduate Center, City University of New York, USA

^b Queens College and the Graduate Center, City University of New York, USA

^c University of Valencia, Spain

^d University of Jaen, Spain

^e University of Otago, New Zealand

ABSTRACT

In this experimental study we investigated the influence of *discrepant* feedback from teachers and peers on emotional responses and feedback appraisals in a sample of university students. A total of 753 tertiary education students from the United States ($N = 172$), New Zealand ($N = 217$), and Spain ($N = 364$), were presented with a scenario wherein they received two (discrepant) feedback messages that varied in terms of their focus and tone (suggestive/neutral vs. evaluative/positive). In the two conditions, the source of feedback was also manipulated. In condition one, participants saw that the teacher offered evaluative/positive feedback, with peers offering suggestive/neutral. In contrast, in condition two, the peer offered evaluative/positive message while the teacher provided suggestive/neutral. The findings from repeated measures ANOVA, cumulative models, and logistic regression revealed a clear preference among students for feedback from teachers over peers, regardless of the message's focus and valence. Positive evaluative messages were found to elicit higher positive emotions. Interestingly, even when peer feedback was viewed as having advantages, students reported higher positive emotions and were more receptive of teacher feedback. These results were consistent across countries, although our findings did reveal country-specific patterns. These insights have practical implications for targeted training on feedback provision, highlighting the pedagogical value of peer feedback and offering important insights to engage students with diverse feedback sources.

1. Introduction

Educational research consistently highlights the potential positive impact of feedback on learning outcomes (e.g., Hattie & Timperley, 2007; Lipnevich & Panadero, 2021; Mandouit & Hattie, 2023; Morris, Perry, & Wardle, 2021; Narciss, Prescher, Khalifah, & Körndle, 2022; Nicol, 2021; Wisniewski, Zierer, & Hattie, 2020). However, its effectiveness varies based on specific characteristics of feedback messages, the instructional context, tasks, and students' and providers' attributes (Lipnevich & Panadero, 2021; Lipnevich & Smith, 2022; Wisniewski et al., 2020). Recent research in education has examined the

effectiveness of feedback across various message characteristics, such as content (Narciss et al., 2014), source (Tomazin et al., 2023), emotional valence (Lipnevich & Smith, 2009b), mode of delivery (Máñez et al., 2024), and timing (Metcalf et al., 2009; Nakata, 2015). Studies have also explored students' needs, perceptions, and preferences for feedback (Barnard et al., 2015; Paterson et al., 2020; Van Boekel et al., 2023). Additionally, there is growing attention to research on the range of emotional responses that can influence students' uptake of instructional feedback (Lipnevich, Murano, Krannich, & Goetz, 2021; Pekrun, Cusack, Murayama, Elliot, & Thomas, 2014).

Despite significant advancements in the field, the affective,

[☆] This article is part of a special issue entitled: 'Advancing Feedback Research' published in Contemporary Educational Psychology.

* Corresponding author at: Department of Educational Psychology, the Graduate Center, The City University of New York, 365 5th Avenue, New York, NY 10016, USA.

E-mail address: cloperaoquendo@gradcenter.cuny.edu (C. Lopera-Oquendo).

¹ ORCID: 0000-0002-9355-5843

² ORCID: 0000-0003-0190-8689

³ ORCID: 0000-0002-9753-3246

⁴ ORCID: 0000-0001-8093-1945

⁵ ORCID: 0000-0002-6980-2680

⁶ ORCID: 0000-0003-4481-8982

cognitive, and behavioral processes that underlie the uptake of feedback received from different sources in the academic context (i.e., peer vs. teacher) remain unclear. It is even less clear how students respond and negotiate responses to multiple messages coming from various sources. This cross-national study relied on samples from the United States (US), Spain (SP), and New Zealand (NZ) to examine college students' preferences when confronted with *discrepant* feedback from teacher and peers and how their personality characteristics predicted their emotional responses, perceptions of helpfulness and utility, and intentions to use feedback for task improvement. To this end, we operationally defined *discrepant feedback* as feedback provided for the same work that differs in message content, emotional valence, focus, and the emphasis on next steps for revisions.

1.1. Instructional feedback

Feedback, as conceptualized within the theoretical framework of Lipnevich and Smith (2022), refers to any information provided in response to a person's performance that has the potential to enhance performance on current or future tasks, as well as long-term learning. According to their model, effective feedback requires active processing by the receiver, and the decision to use feedback relies on a complex interplay of characteristics encompassing the context, source, message, and recipient. This theoretical framework is central to this study, providing a lens through which the dynamics of students' appraisal and engagement with feedback are examined (for a review of the Lipnevich & Smith (2022) model, see Fig. S1 in Supplementary Materials).

In academic contexts, students receive feedback in a variety of forms (e.g., comments and grades) and from different sources, such as teachers and peers. Thus, factors such as the source's credibility, the clarity and relevance of the message, and the recipient's individual characteristics all need to align for feedback to be impactful. This alignment ensures that the feedback is not only delivered accurately but is also tailored to the recipient's needs, promoting a more meaningful and constructive engagement (Lipnevich & Smith, 2022; Winstone et al., 2017).

Despite the general agreement that feedback is essential to enhance learning and performance (Lipnevich & Panadero, 2021; Morris et al., 2021; Wisniewski et al., 2020), students often choose to reject feedback (Harris et al., 2014, 2018; Ramani et al., 2019). Researchers have reported several reasons behind students' choice to not use feedback, which include, for example, lack of perceived utility and helpfulness, scarcity of details and individualization, unfriendly tone of feedback provider, lack of respect for the provider, and poor student-teacher relationship (Jonsson, 2013; Vattøy et al., 2021; Weaver, 2006). Conversely, perceived source trust, for example, has shown to affect perceptions of accuracy and overall satisfaction with feedback positively (Rotsaert et al., 2017; Van de Ridder et al., 2014).

As it is also discussed in the Student-Feedback Interaction Model (Lipnevich & Smith, 2022), individual characteristics of the feedback recipient have been shown to affect feedback uptake (Aben et al., 2022; Shute, 2008; Wingate, 2010). For example, Lipnevich and colleagues (Lipnevich et al., 2020; Lipnevich & Lopera-Oquendo, 2022) have shown that variability in students' level of receptivity to feedback predicted their academic performance as well as other meaningful outcomes. Interestingly, students' personality is also related to receptivity, with conscientiousness and openness being the strongest predictors of students' willingness to engage with feedback. Further, learners who present higher levels of self-efficacy, achievement, as well as self-regulation, are more likely to engage with feedback actively (Adams et al., 2020; Baadte & Schnotz, 2014; Gan et al., 2023). In the following sections, we will take a closer look at students' responses to varying feedback messages and the potential influences of different sources on students' performance, learning, and other key variables.

1.2. Feedback message

A key component of the Student-Feedback Interaction Model (Lipnevich & Smith, 2022) is the content of the feedback provided to students, as the message can strongly impact learners' engagement with feedback and, consequently, their learning and performance. Several characteristics of the feedback message, such as level of detail, comprehensibility, function, and tone, have been shown to play a role in how feedback is received by the learner (Lipnevich & Smith, 2022). Students often report their preference for feedback that identifies and explains what needs to be improved (Hattie et al., 2021; Hattie & Timperley, 2007; Lipnevich & Smith, 2009b). On the other hand, praise (along with criticism) is considered an ineffective form of feedback (Hattie & Zierer, 2019). Feedback that is rich in praise tends to hinder students' motivation (Hill et al., 2021; Lipnevich et al., 2023; Lipnevich & Smith, 2009a) as well as their improvement on a task (Hattie et al., 2021), although teacher praise has been found an effective tool to manage students' behavior in class (Floress et al., 2017).

Feedback plays a crucial role in student performance. Hattie et al. (2021) found that "Where to Next" feedback, which guides future improvements, enhances writing outcomes, with the number of comments being the strongest predictor of progress. Similarly, Agius and Wilkinson's (2014) review showed that students value feedback that offers directions for enhancement, a trend also evident in peer feedback research. Constructive and cognitive-focused peer feedback has been linked to greater improvements than affective feedback, such as praise (Cheng et al., 2015; Kerman et al., 2022). The impact of praise on student's motivation is complex. Henderlong & Lepper (2002) emphasized that its effects depend on perceived sincerity, task difficulty, and attributions. Praise fostering autonomy and competence can enhance motivation, yet Lipnevich et al. (2023) caution against praise overshadowing critical feedback, leading to decreased motivation and performance. Considering these insights, our study will consider two types of messages: a suggestive, "Where to next" type of feedback (Hattie et al., 2021) and overall positive, evaluative feedback, or praise.

1.3. Sources of feedback

While feedback can come from different sources (Lipnevich & Smith, 2022), students often prefer teacher feedback over feedback from their peers (Ekoniak & Paretti, 2018; Mahvelati, 2021; Van Der Kleij & Lipnevich, 2021; Yang, Badger, & Yu, 2006; Zacharias, 2007). A possible explanation for that is that students' perceptions of the credibility of the feedback provider may influence how they use certain feedback (Albright & Levy, 1995; Jonsson & Panadero, 2018; Rotsaert, Panadero, Estrada, & Schellens, 2017; Ruegg, 2015). Conversely, peer feedback is often met with skepticism due to concerns about trustworthiness and accuracy (Dijks, Brummer, & Kostons, 2018; Rotsaert, Panadero, Estrada, & Schellens, 2017). Nevertheless, research has demonstrated the positive effects of both teacher and peer feedback on student performance (Graham, Hebert, & Harris, 2015; Huisman, Saab, Van Den Broek, & Van Driel, 2019; Lv, Ren, & Xie, 2021; Sun & Wang, 2022). One possible explanation for this is that, despite concerns about credibility, peers generally provide positive and supportive comments, which are perceived as beneficial. Therefore, the value of effective feedback mitigates doubts about peer competence, objectivity, and potential negative interpersonal dynamics, ultimately fostering engagement and learning.

If feedback from only one source was not sufficiently complex, students' responses to discrepant messages from multiple sources elevate the complexity to an entirely new level. Albright and Levy (1995) examined how undergraduate psychology students reacted to feedback from a peer and professor feedback, considering source credibility and rating discrepancies. In a lab study, students engaged in a managerial problem-solving task, self-assessed their performance, and then received fictitious evaluations, one aligned with their self-assessment and the other differing in positivity or negativity. Instructor feedback was rated

as more useful and accurate than peer feedback. Positive discrepancies led to favorable ratings, whereas negative discrepancies resulted in lower ratings. To our knowledge, this is the only study that investigated discrepant feedback from different sources, but the nature of the task was such that participants' responses were guided by professional values and beliefs. Our study, on the other hand, concerns discrepant instructional feedback from a peer and a teacher occurring in an academic context and serves as the initial step in examining this very common yet under-investigated phenomenon.

1.4. Cross-National experiences with instructional feedback

Drawing upon the findings of the PISA 2018 report (OECD, 2019), it is evident that while students generally feel supported by their teachers, only a small fraction, about 10 to 15 %, reported receiving feedback regularly during lessons. Alarmingly, a significant portion of students mentioned that they rarely or never receive any feedback at all. Even among OECD countries, robust differences can be observed in students' achievement and educational experiences. In our study we examined student responses to feedback in the United States, Spain, and New Zealand, thus covering three continents and three distinct educational systems. According to the PISA 2022 (OECD, 2023) assessments, students' performances in mathematics presented some variability (US: $M = 465$, $SD = 95$; NZ: $M = 479$, $SD = 99$; SP: $M = 473$, $SD = 86$), but that variability was even more significant in the context of reading (US: $M = 504$, $SD = 111$; NZ: $M = 501$, $SD = 109$; SP: $M = 474$, $SD = 97$). Considering assessment practices, a significant portion of teachers from these three countries reported actively engaging in assessing their student's progress and providing feedback, all exceeding the OECD average of 79 %. In the United States, 85 % of teachers reported routinely observing their students and offering immediate feedback, whereas in New Zealand, the proportion was 89 %, and in Spain, it was 83 % (TALIS, 2024).

Regarding feedback practices in New Zealand schools, Harris et al. (2014) reported that teachers in elementary and secondary schools employed formal written reports and ongoing formative feedback during instruction. Moreover, Brown et al. (2012) identified that practicing teachers highlighted an emphasis on learning-oriented feedback rather than using feedback primarily to bolster student well-being, such as praising effort or boosting self-esteem. In the United States, students generally enjoy receiving feedback. Zumbrunn et al. (2016) showed that 80% of participants in their study, middle and high school students affirmed their appreciation for feedback on their writing assignments. Feedback practices in Spain rely heavily on the teacher, with Quesada Serra et al. (2016) reporting that professors valuee monitoring student learning and feel adequately capable of carrying out assessment tasks for that purpose. Although the studies involve samples ranging from primary to college level students, they provide foundational insights into feedback practices that can be relevant across educational levels. Despite certain similarities, the three countries represent a unique context wherein we will consider students' responses to discrepant messages.

1.5. The current study

To investigate students' preferences for feedback that guide their intention to use or reject specific feedback, we employed a mixed-factor design to examine how variations in feedback content and tone, delivered by two distinct sources, shape students' perceptions and reactions. The study utilized vignettes that contextualized the task and manipulated both the message type (evaluative/positive vs. suggestive/neutral) and the source of feedback (teacher or peer). Furthermore, we assessed personality traits and individual characteristics (e.g., gender, age, and GPA) to identify factors that predict college students' responses to discrepant feedback from teachers and peers. We were also interested in examining cross-national patterns. Thus, we included samples of college

students from United States, Spain, and New Zealand. These countries represent diverse educational, cultural, and policy contexts, making their inclusion highly valuable for examining global trends in education. Spain offers insights into systems influenced by European Union policies and a strong emphasis on regional autonomy, including variations in language and curriculum. New Zealand provides a unique perspective as a small Pacific nation with a bicultural framework grounded in the Treaty of Waitangi, highlighting equity for Māori and Pasifika students and a competency-driven national curriculum. The United States, with its size and decentralized educational system, reflects a mix of federal oversight and local control, characterized by significant socioeconomic and cultural diversity. Thus, the inclusion of three countries enables a meaningful comparative exploration (Marín et al., 2022; Tan & Pillay, 2008).

In this study, we aimed to address the following research questions:

- 1. Are there differences in students' perceptions of utility, helpfulness, and intentions to use discrepant instructional feedback when it comes from different sources?
- 2. Are there differences in students' reported emotions towards discrepant feedback messages depending on its source?
- 3. Do students' personality traits and receptivity to feedback predict their responses to *discrepant* peer and teacher feedback?
- 4. To what extent can these findings be replicated across three different contexts of the United States, New Zealand, and Spain?

2. Method

2.1. Participants

A total of 753 tertiary education students from the United States ($N = 172$), New Zealand ($N = 217$), and Spain ($N = 364$) participated in this study. Participants who omitted more than 80% of items in the instrument were deleted from the final dataset, so 735 observations were retained. 73% of participants were female and 88% were in the age range between 18 and 25 years old ($M = 21.5$, $SD = 5.35$), and 91% were enrolled in an undergraduate program. The demographic characteristics of the participants varied across the three countries (Table 1).

In terms of sex at birth, the majority of participants were female in Spain (84%), followed by the United States (77%), and New Zealand (53%). Regarding age distribution, participants under 25 constituted 95% and 94% in New Zealand and Spain, respectively, whereas, in the United States, 75% fell into this age bracket. Furthermore, educational levels varied, with undergraduate students comprising 68% in the United States, 95% in New Zealand, and 100% in Spain. Graduate-level participants coming from the first year of a teacher preparation master program were represented by 32% in the United States and 5% in New Zealand.

Table 1
Demographic Statistics by Country (N = 735).

Demographic Variable	United States (N = 171)	New Zealand (N = 206)	Spain (N = 358)
Sex at birth			
Female	77% (132)	53% (109)	84% (299)
Male	22% (37)	45% (95)	16% (59)
Age			
< 20 years	29% (49)	88% (181)	53% (190)
21–25 years	47% (80)	7% (15)	40% (144)
> 26 years	25% (42)	5% (10)	6% (23)
Educational Level			
Undergraduate	68% (117)	95% (196)	100% (358)
Graduate	32% (54)	5% (10)	

Note: n participants by group in parenthesis.

2.2. Procedure

The study was conducted by distributing an online survey using the Qualtrics platform. Feedback messages were originally created in Spanish and translated. We used a double translation procedure to produce the English version of the comments. First, the participants were asked to provide demographic information and academic background and respond to the Big Five personality inventory (Benet & John, 1998). Next, they were asked to imagine a scenario wherein they received (discrepant) feedback from their teacher and peers. Therefore, they were given the following instructions:

“Imagine you have just finished a practice presentation for your upcoming college class project. You put in a lot of effort to make it the best possible. You know you will have to present it to a larger audience next week and want to ensure it is perfect. Below are some feedback comments about your presentation provided for both your course instructor, whom you respect, and one of your smartest classmates. Please read these messages carefully. In the next sections, you will find questions about your perceptions of these comments.”

All participants were randomly assigned to one of two feedback scenario conditions. Randomization was implemented through Qualtrics, utilizing a randomized, evenly distributed approach that ensured each condition was presented an equal number of times across administrations. In both conditions, students simultaneously received two different feedback messages that varied in function and tone. The feedback was manipulated as follows:

- Comment 1 (Evaluative and positively valenced): included a positive summary of performance that did not include any future-oriented suggestions for improvement.
- Comment 2 (Suggestive and neutrally valenced): included a range of suggestions for future improvement delivered in a neutral tone.

In each condition, the source of feedback was also manipulated, so in condition one, the participants saw that the teacher offered evaluative/positive feedback, and the peer offered suggestive/neutral. In contrast, the peer offered an evaluative/positive message in condition two, while the teacher provided a suggestive/neutral one (see Table 2).

In summary, the study employed a 2x2 mixed-factor design, with the type of message serving as a between-subjects factor and the source of feedback (i.e., teacher vs. peer) as a within-participant factor. The study involved random assignment of participants to different feedback conditions, which is a core characteristic of experimental design (Bordens et al., 2014; Cohen et al., 2007). Tables D1 and D2 (Appendix D, Supplementary Materials) present the messages provided to participants in English and Spanish version. The messages were virtually identical in content and structure, with only slight adjustments made to reflect the natural differences in communication style between peers and teachers. These subtle differences were intentionally included to enhance the realism of feedback exchanges, where feedback from peers and teachers often varies in formality and tone. This way, the ecological validity of the intervention was bolstered.

After reading the two discrepant messages, the participants were asked to review each message (teacher and peer) and rate 1) the emotions that each message would elicit, and 2) the utility, helpfulness, and intention to incorporate the feedback message (feedback appraisal). Later, they were asked to rank which comment (teacher or peer) would

provide more valuable feedback to improve their final presentation. Finally, participants filled the global feedback receptivity scale (Lipnevich & Lopera-Oquendo, 2024; Lipnevich, Murano, Krannich, & Goetz, 2021)

Participants received invitations from their instructors to take part in the study and provided their consent to participate. Although participation was voluntary, certain instructors framed the activity as an optional extra credit opportunity, providing students with additional credit for their involvement. The university’s ethics committee approved the study protocol (2022–0460-QC).

2.3. Instrumentation

Big Five Personality Inventory (BFI). The BFI is a 44-item inventory that measures extraversion, agreeableness, conscientiousness, neuroticism, and openness (Benet & John, 1998). Responses to each personality indicator ranged from 1 (strongly disagree) to 5 (strongly agree). Composite scores were derived by averaging the responses corresponding to the five-personality trait. The internal consistency reliability statistics (Cronbach’s α) across the five scales ranged from $.71 < \alpha < .85$ in the United States, $.60 < \alpha < .83$ in New Zealand, and $.70 < \alpha < .84$ in Spain (see Table S1 Supplementary Materials).

Emotions. Eight single items based on a brief measure of Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) were used to gauge discrete emotions that teacher and peer feedback messages could elicit (i.e., distressed, upset, enthusiastic, inspired, scared, irritable, proud, and happy). The items prompted participants to report how they would feel about the feedback their teacher or peers provided, using a Likert scale ranging from 1 (very slightly or not at all) to 5 (extremely). Individual scores for discrete emotions were used in the current analysis.

Feedback appraisal. Three single items were used to assess participants’ perception of each type of comment (teacher and peer) to provide improvement insights. Specifically, two items assessed the level of participants’ agreement, using a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree) regarding the perceived utility (“The information provided in the comment would be useful to me”) and helpfulness (“The information provided in the comment could help me improve the presentation”) of each feedback message. An additional item measured the intention to incorporate each feedback message using a 5-point scale (1 = Extremely unlikely and 5 = Extremely likely) (“Given the scenario presented above to what extent would you integrate the feedback provided by your peer/teacher to improve your final presentation?”). Individual discrete items were used in the current analysis.

Overall message value. A single item evaluated the overall message value derived from both sources. Participants were asked to rank (1 = most valuable, 0 = less valuable) the feedback message they perceived as most valuable for improving their presentation (“Based on the scenario described earlier, which feedback do you find the most valuable in enhancing your final presentation?”).

Global Receptivity to Feedback. The Global Receptivity to Feedback scale is a new self-report instrument designed to assess general attitudes and dispositions to receiving external feedback or comments about progress or performance across domains (Lipnevich & Smith, 2009; Lipnevich et al., 2016). This new construct includes 8 Likert-type items on a 5-point scale (1 = strongly disagree and 5 = strongly agree). A unidimensional CFA model was used to extract individual scores. The internal consistency reliability statistics (Cronbach’s α) were .81, .79, and .73 for the United States, New Zealand, and Spain, respectively (see Tables S1 Supplemental Material).

Demographic and academic information. Information about students’ sex at birth, age, and GPA were also collected.

2.4. Analytic Plan

Descriptive and psychometric analyses were conducted initially to evaluate the reliability and validity of multi-item scale measures. First,

Table 2
Description of type of condition.

Condition	Type of message by source	
	Teacher	Peer
Condition 1	Evaluative/positive valence	Suggestive/neutral valence
Condition 2	Suggestive/neutral valence	Evaluative/positive valence

exploratory Factor Analysis (EFA), using parallel analysis and oblimin rotation of polychronic correlation matrix, as well as Confirmatory Factor Analysis (CFA) through Structural Equation Modeling (SEM) with weighted least squares mean and the variance adjusted (WLSMV) estimator (Sass et al., 2014) was conducted to examine the factor structure of global receptivity to feedback scale. We used the following indices to assess the overall fit model good fit: (1) the Root Mean Square Error of Approximation (RMSEA) with values < 0.08 being indicative of reasonable fit and values < .05 indicating a good fit; (2) the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) with values > .90 indicating an acceptable fit and values > .95 indicating a good fit; and (3) the Standardized Root Mean Square Residual (SRMR) with values < .05 being indicative of good fit (Brown, Harris, & Harnett, 2014; Browne & Cudeck, 1992; Cheung & Rensvold, 2002; Hair, Black, Babin, & Anderson, 2010; Hu & Bentler, 1998, 1999; MacCallum, Browne, Sugawara, & Wegener, 1996; Yu, 2002).

Second, we conducted a set of nested hierarchical models to test hypotheses about CFA data fit, including the configural, metric, scalar, partial scalar, and partial strict factorial invariance models, as well as equal means, variance and covariance factor models (Byrne, Shavelson, & Muthén, 1989; Gregorich, 2006; Lugtig, Boeije, & Lensvelt-Mulders, 2012; Millsap & Yun-Tein, 2004; Schoot, Lugtig, & Hox, 2012; Vandenberg & Lance, 2000; Milfont and Fischer, 2010), for addressing the cross-country invariance of personality traits and global receptivity to feedback scale. Measurement invariance was tested by evaluating how well the models fit the observed data. Configural invariance was tested by evaluating the overall fit of the models. Moreover, the comparison of nested models, that is, models that are identical except for a target set of restrictions in one, was evaluated by comparing the differences in fit statistics for the two models. Specifically, Satorra-Bentler (SB) scaled statistics ($\Delta\chi^2$) and differences in model degrees of freedom (Δdf) were computed to test whether the more constrained model resulted in a significantly weaker fit. We considered a violation of invariance when comparing one level of analysis to the next more stringent level, we found a change in CFI greater than .01 together with a change in RMSEA greater than .015 (Chen, 2007; French & Finch, 2006; Putnick & Bornstein, 2016).

To address the first and second research questions aimed to compare differences in emotional responses and feedback appraisals elicited by discrepant messages, a repeated measures ANOVA was conducted, with conditions and country as between factors and source of feedback (i.e., teacher vs. peer) as a within-participant factor. To answer the third and fourth research question, we followed two approaches. First, we conducted a cumulative or graded response model (Bürkner & Vuorre, 2019; Samejima, 1969), with a logit function link, to predict the effects of type of feedback message (i.e., evaluative/praise vs suggestive/neutral) and individual characteristics (i.e., global receptivity to feedback and personality) on ordinal responses to elicited emotions and appraisal for each source of feedback (i.e., teachers and peer) by country. Second, we conducted a logistic regression model with a logit link function to estimate the effect of the experimental condition, global receptivity to feedback, and personality on the probability of ranking the teacher comments as the most valuable feedback message. All analyses were carried out using R software version 4.1.2 (R Core Team, 2024; Rosseel, 2012). To control for the inflation of Type I error, p-values were adjusted using the Benjamini–Hochberg correction (Benjamini & Hochberg, 1995; Feise, 2002). The fourth research question, concerning the cross-country comparison of results across the United States, New Zealand, and Spain, has been addressed through the analyses conducted in the preceding research questions.

3. Results

3.1. Preliminary analysis

A Confirmatory Factor Analysis (CFA) through Structural Equation

Modeling (SEM) with a weighted least square mean and variance adjusted (WLSMV) estimator (Sass et al., 2014) was conducted to examine the unidimensional structure of the Global Receptivity Feedback scale. After adding correlation between items suggested by modification indices, the measure model showed a good fit to the data: $\chi^2(47) = 45.16, p < .001$; RMSEA = .06 (90% CI: .04, .08), CFI = .995, TLI = .989, and SRMR = .039. Table A1 (Appendix A, Supplementary Materials) shows all original items of the scale and their final factor loading. Invariance analysis of the measure across the three countries of administration was also conducted. Findings supported measurement equivalence (i.e., configural, metric, and partial strict) for the global feedback scale by the country (Appendix A, Table A2 Supplementary Material). Latent factor scores, with a mean of 0 and a standard deviation of 1 for the full sample, were saved and used in the cumulative and logistic models.

The Big-5 Inventory (BFI) is a widely used tool for assessing personality traits across five dimensions. The BFI generally shows good internal consistency, with Cronbach's alpha values typically ranging from .70 to .80 across different domains. The factor structure of the BFI is consistently replicated across different languages and cultural contexts (Gallardo-Pujol et al., 2022; Thalmayer & Saucier, 2014) while providing evidence of structural validity. Confirmatory Factor Analysis was conducted to check the structural validity of personality traits across countries in our study. Findings supported configural measurement equivalence of the five-factor multi-correlated model by countries, which showed a similar fit to the full sample (See Appendix A, Tables A3, A4, and A5, Supplementary Material). However, metric invariance was not supported based on the Satorra-Bentler (SB) scaled chi-square difference test ($\Delta\chi^2(78) = 137.04, p < .001$). This result suggests that constraining factor loadings to equality across countries resulted in a significant worsening of model fit, indicating that the contribution of individual items to the latent constructs differs across countries.

3.2. Descriptive analysis

Table S2 (Supplementary Materials) provides overall descriptive statistics for participants' emotional responses and feedback appraisals by country. Across all participants, the highest rating for negative emotions elicited by teacher and peer feedback messages was distress, with 24.1% ($M = 1.86, SD = 1.07$) of participants indicating they would feel this emotion (i.e., moderately, quite a bit, or extremely) after receiving teacher feedback message. Similarly, 28.7% ($M = 1.93, SD = 1.17$) of participants expressed feeling distressed after receiving feedback from peers in both conditions. In terms of positive emotions, the highest ratings were for feelings of pride and happiness, with 54.1% ($M = 3.46, SD = 1.45$) and 47.5% ($M = 3.20, SD = 1.44$) of participants indicating they would feel these emotions after receiving feedback from teachers and peers, respectively.

Regarding feedback appraisals, 86% ($M = 4.22, SD = 0.86$) and 72% ($M = 3.91, SD = 1.22$) of participants agreed or strongly agreed that the information provided in teacher comments would be helpful and useful. Additionally, 73% ($M = 4.05, SD = 1.186$) indicated they were likely to incorporate this feedback to enhance their task. In contrast, 63.9% ($M = 3.82, SD = 0.99$), 69.7% ($M = 3.667, SD = 1.19$), and 60.7% ($M = 3.725, SD = 1.11$) of participants expressed similar opinions regarding the helpfulness, usefulness, and intention to use peer feedback comments, respectively. Table S3 (Supplementary Materials) presents trends by country and condition, while Tables S4 and S5 show the standardized and unstandardized scores for individual characteristics (i.e., personality and global feedback) by country. Additionally, Appendix B includes descriptive information by gender and education level.

3.3. Differences in students' feedback appraisals of discrepant peer and teacher feedback

In order to examine the first research question, a repeated-measured

ANOVA with the condition and country as between factors and source of feedback (i.e., teacher vs. peer) as a within-participant factor was conducted to determine whether or not there were differences in means for feedback appraisal items (Table 3).⁷ The results showed a statistically significant main effect of country and source on means for feedback appraisal items, as well as a significant interaction effect of condition and source for helpfulness ($F(1, 715) = 139.831, p < .001$), usefulness ($F(1, 715) = 507.757, p < .001$), and intention to use feedback ($F(1, 715) = 193.930, p < .001$). Moreover, a significant interaction effect was observed between conditions, source, and country, indicating that the impact of condition and source on feedback appraisal varied depending on country (Helpfulness: $F(2, 715) = 22.290, p < .001$, Usefulness: $F(2, 715) = 12.795, p < .001$, and intended to use feedback: $F(2, 715) = 12.879, p < .001$).

Results of multiple comparisons of rating means (Table S6, Supplementary Material) within conditions (i.e., teacher suggestive compared to teacher evaluative) revealed that participants rated suggestive/neutral feedback higher on usefulness, helpfulness, and intent to use information, compared to the evaluative/positive one. Specifically, in Condition 1, feedback appraisals elicited by peer suggestive/neutral messages were significantly higher than those elicited by teacher evaluative/positive comments. Similarly, in Condition 2, feedback appraisals elicited by teacher suggestive/neutral comments were significantly higher than those elicited by peer evaluative/positive. These results were consistent across countries.

Comparisons between suggestive/neutral versus evaluative/positive

Table 3
Repeated Measure ANOVA for Feedback Appraisal within-subjects Factor.

Effect	DFn	DFd	F	p		η^2
Helpfulness						
Condition	1	715	2.147	.143		0.003
Country	2	715	27.997	.000	*	0.073
Source	1	715	55.644	.000	*	0.072
Condition*Country	2	715	2.267	.104		0.006
Condition*Source	1	715	139.831	.000	*	0.164
Country*Source	2	715	2.806	.061		0.008
Condition*Country*Source	2	715	22.290	.000	*	0.059
Usefulness						
Condition	1	715	2.164	.142		0.003
Country	2	715	27.648	.000	*	0.072
Source	1	715	6.750	.010	*	0.009
Condition*Country	2	715	2.254	.106		0.006
Condition*Source	1	715	507.757	.000	*	0.415
Country*Source	2	715	1.867	.155		0.005
Condition*Country*Source	2	715	12.795	.000	*	0.035
Intent to Use						
Condition	1	715	2.527	.112		0.004
Country	2	715	7.943	.000	*	0.022
Source	1	715	23.217	.000	*	0.031
Condition*Country	2	715	0.656	.519		0.002
Condition*Source	1	715	193.930	.000	*	0.213
Country*Source	2	715	0.379	.685		0.001
Condition*Country*Source	2	715	12.879	.000	*	0.035

Note: η^2 (generalized η^2) represents the effect size is essentially the amount of variability due to the within-subjects factor ignoring the effect of the subjects. Huynh-Feldt sphericity correction (Girden, 1992) was applied for all within-subjects factors even if the assumption of sphericity was met, $\epsilon > .75$ in all cases when sphericity correction was not met. * $p < .05$.

⁷ The sample size was determined through power analysis conducted in pwrss R package (Bulus, M, 2023) using a specification of a repeated measure model with an interaction effect model. Expecting a difference of Cohen's $d = 0.30$ and a correlation of 0.50 between teacher and peer ratings ($n.rm = 2$) after controlling for the group, translating into partial $\eta^2 = 0.020$ with condition and country groups (levels = 6), the recommended sample size of 324. Therefore, Statistical power = 0.8, Type I error rate = 0.05 and Type II error rate = 0.2.

feedback provided by the same source revealed that the means of ratings for feedback appraisal items were significantly higher for suggestive/neutral comments than for evaluative/positive ones, independently of the source (i.e., teacher or peer). Results also showed that participants demonstrated significantly higher appraisal for teacher messages compared to those from peers, with some differences across countries. For example, in New Zealand, there were no significant differences in the mean ratings for usefulness ($p = .591$) and helpfulness ($p = .857$), while in the United States, no differences were observed in means for usefulness ($p = .205$) between suggestive/neutral messages provided by peers and teachers. Regarding evaluative/positive comments, no significant differences existed between means of intention to use across countries.

3.4. Differences in students' emotional responses to discrepant peer and teacher feedback

To answer the second research question, a repeated-measures ANOVA with conditions and country as between-subject factors and the source of feedback (i.e., teacher vs. peer) as a within-participant factor was conducted to determine differences in negative and positive emotions elicited by feedback messages (Table 4). The results revealed a statistically significant main effect of condition, country and source on means for all negative discrete emotions, as well as a significant interaction effect of condition and source: distressed ($F(1, 716) = 563.659, p < .001$), upset ($F(1, 715) = 499.915, p < .001$), scared ($F(1, 714) = 262.562, p < .001$) and irritable ($F(1, 715) = 376.984, p < .001$). Moreover, a significant interaction effect was observed among conditions, source, and country for distressed ($F(2, 715) = 3.271, p < .039$) and scared ($F(2, 715) = 14.094, p < .001$), indicating that the impact of condition and source on these negative emotions also varied depending on country.

For positive emotions, a repeated-measures ANOVA indicated statistically significant main effects of country and source on the means for all positive discrete emotions. A significant interaction effect between condition and source was also observed for enthusiastic ($F(1, 714) = 935.692, p < .001$), inspired ($F(1, 713) = 261.433, p = .006$), happy ($F(1, 713) = 1356.775, p = 0.003$) and proud ($F(1, 713) = 1443.187, p < .001$).

Furthermore, a significant three-way interaction effect among condition, source, and country was identified for enthusiastic ($F(2, 714) = 12.517, p < .001$), inspired ($F(2, 713) = 19.656, p < .001$), and proud ($F(2, 713) = 5.750, p = .003$), suggesting that the effect of condition and source on these positive emotions also varied by country. Results of the multiple comparison tests are provided in Tables S6 and S7 (Supplementary Materials). Overall, the findings suggested that positive emotions elicited by evaluative/positive messages were significantly higher ($p < .05$) than those elicited by suggestive/neutral messages, regardless of the source. Conversely, negative emotions elicited by suggestive/neutral messages were significantly higher ($p < .05$) than those elicited by evaluative/positive messages.

3.5. Individual characteristics as predictors of students' responses to peer and teacher feedback

A series of cumulative models (also known as graded response) (Bürkner & Vuorre, 2019; Samejima, 1997), with a logit function link was used to predict the effects of type of feedback message (i.e., evaluative/praise vs suggestive/neutral) and individual characteristics (i.e., age, sex at birth, GPA, global receptivity to feedback and personality) on ordinal responses to elicited emotions and feedback appraisal. Models were estimated by the source of feedback (i.e., teachers and peers) and country independently, and then a total of two models were conducted by country. All variables were standardized by country. Results for cumulative models are displayed in Tables 9 and 10. To control for the inflation of Type I error, p-values were adjusted using the

Table 4

Repeated Measure ANOVA for Emotional response a within-subjects factor.

Effect	Positive Emotions						Negative Emotions							
	Variable	DFn	DFd	F	p	η^2	Variable	DFn	DFd	F	p	η^2		
Condition	Enthusiastic	1	714	0.646	.422	0.001	Distressed	1	716	6.053	.014	*	0.008	
Country		2	714	5.469	.004	*		0.015	2	716	0.873	.418		0.002
Source		1	714	17.301	.000	*		0.024	1	716	3.906	.048	*	0.005
Condition*Country		2	714	1.674	.188			0.005	2	716	2.199	.112		0.006
Condition*Source		1	714	935.692	.000	*		0.567	1	716	563.659	.000	*	0.440
Country*Source		2	714	11.037	.000	*		0.030	2	716	1.814	.164		0.005
Condition*Country*Source	2	714	12.517	.000	*	0.034	2	716	3.271	.039	*	0.009		
Condition	Inspired	1	713	0.009	.923	0.000	Scared	1	714	3.944	.047	*	0.005	
Country		2	713	5.261	.005	*		0.015	2	714	7.401	.001	*	0.020
Source		1	713	24.751	.000	*		0.034	1	714	9.538	.002	*	0.013
Condition*Country		2	713	0.725	.485			0.002	2	714	3.828	.022	*	0.011
Condition*Source		1	713	261.433	.000	*		0.268	1	714	262.562	.000	*	0.269
Country*Source		2	713	5.145	.006	*		0.014	2	714	0.575	.563		0.002
Condition*Country*Source	2	713	19.656	.000	*	0.052	2	714	14.094	.000	*	0.038		
Condition	Happy	1	713	0.101	.750	0.000	Irritable	1	715	6.194	.013	*	0.009	
Country		2	713	15.205	.000	*		0.041	2	715	1.424	.241		0.004
Source		1	713	26.138	.000	*		0.035	1	715	24.846	.000	*	0.034
Condition*Country		2	713	2.035	.131			0.006	2	715	3.608	.028	*	0.010
Condition*Source		1	713	1356.775	.000	*		0.656	1	715	376.984	.000	*	0.345
Country*Source		2	713	5.967	.003	*		0.016	2	715	5.349	.005	*	0.015
Condition*Country*Source	2	713	5.750	.003	*	0.016	2	715	1.864	.156		0.005		
Condition	Proud	1	714	0.993	.319	0.001	Upset	1	715	7.353	.007	*	0.010	
Country		2	714	30.198	.000	*		0.078	2	715	1.251	.287		0.003
Source		1	714	52.627	.000	*		0.069	1	715	22.918	.000	*	0.031
Condition*Country		2	714	3.626	.027	*		0.010	2	715	10.012	.000	*	0.027
Condition*Source		1	714	1443.187	.000	*		0.669	1	715	499.915	.000	*	0.411
Country*Source		2	714	8.885	.000	*		0.024	2	715	11.284	.000	*	0.031
Condition*Country*Source	2	714	1.709	.182		0.005	2	715	1.906	.149		0.005		

Note: η^2 (generalized η^2) represents the effect size is essentially the amount of variability due to the within-subjects factor ignoring the effect of the subjects. Huynh-Feldt sphericity correction (Girden, 1992) was applied for all within-subjects factors even if the assumption of sphericity was met, $\epsilon > .75$ in all cases when sphericity correction was not met. * $p < .05$.

Benjamini–Hochberg correction (Benjamini & Hochberg, 1995; Feise, 2002).⁸

The findings for teacher feedback models (Table 5) indicated that receiving suggestive/neutral comments significantly increased the likelihood of expressing agreement (i.e., agree or strongly agree versus strongly disagree) regarding the helpfulness and usefulness of the teacher's message. Moreover, participants who received such comments were more inclined (i.e., likely or extremely likely versus unlikely) to report higher intention to utilize this information for enhancing their tasks compared to those who received evaluative/positive messages while holding all other variables constant. These results remain consistent across different countries. In the United States, for participants who received suggestive/neutral comments, the odds ratios of showing a stronger agreement that the message was helpful and useful were 4.036 ($\beta = 1.395$, $p < .001$) and 12.86 ($\beta = 2.554$, $p < .001$) times higher, respectively, compared to those who received evaluative/positive messages. Regarding the intent to use the message, the odds ratio was 7.02 ($\beta = 1.95$, $p < .001$) times higher.

Similarly, in New Zealand, these odds ratios were notably higher at 5.68 ($\beta = 1.737$, $p < .001$), 28.93 ($\beta = 3.340$, $p < .001$), and 12.8 ($\beta = 2.550$, $p < .001$) times, respectively, for the same comparisons. In Spain, there were no significant effects of the type of message on the appraisal of the helpfulness of the teacher message ($\beta = 0.373$, $p = .203$), while the odds for usefulness and intent to use suggestive/neutral were 7.88 ($\beta =$

2.064, $p < .001$) and 2.98 ($\beta = 1.091$, $p < .001$) times higher than evaluative/positive teacher' comment. When examining individual characteristics diverse patterns emerged across different countries. The findings indicated that agreeableness had a significant and positive effect on the perceived helpfulness of teacher comments in the United States ($\beta = 0.707$, $p < .001$).

Conscientiousness showed negative effects on the usefulness ($\beta = -0.575$, $p < .010$) of the teacher's feedback for the New Zealand sample. Specifically, for every standard deviation unit increase in conscientiousness traits, the probability of selecting the highest level of agreement (i.e., agree or strongly agree versus strongly disagree) that the teacher's message was useful decreased by 43.7%. The results also revealed a significant and positive effect of global receptivity on the appraisal of teacher feedback in Spain. According to the results, for each standard deviation unit increase in global feedback receptivity, there was an increase in the likelihood of agreeing that the teacher's message was helpful and useful and in the probability of intending to use the feedback by 63.5% ($\beta = 0.492$, $p < .001$), 68.7% ($\beta = 0.523$, $p < .001$), and 54.8% ($\beta = 0.437$, $p = .002$), respectively. Moreover, moderate exposure to teacher feedback resulted in a reduction of 58% in the probability of perceiving teacher comments as helpful ($\beta = -0.892$, $p = .020$) in New Zealand. There were no significant effects of age, sex at birth, and GPA on the probability of responses to feedback appraisal due to teacher messages.

Results regarding peer feedback (Table 6) mirrored those of teacher feedback, showing a significant increase in agreement regarding the helpfulness and usefulness of the peer's message when receiving suggestive/neutral comments. Furthermore, participants who received such comments were more prone to utilize comments for task enhancement compared to those receiving evaluative/positive messages, with all other variables held constant. In the United States, participants receiving suggestive/neutral comments had odds ratios of 3.36 ($\beta = 1.212$, $p < .001$) and 7.19 ($\beta = 1.972$, $p < .001$) times for agreement on helpfulness and usefulness of peer comments, respectively, compared to

⁸ The sample size was determined through power analysis conducted in pwrss R package (Bulus, M, 2023) using probability specification. Assuming a squared multiple correlation of 0.15 between covariables (R-square), a base probability $P_0 = 0.2$, which is the overall probability of being in group 1 without influence of predictors in the model (null), a $P_1 = 0.1$, which is the probability of being in group 1 (P1) deviate from P_0 depending on the value of the predictor under alternative hypothesis, a power of 0.80, and a significance level of 0.05, the recommended sample size for a logistic regression is 143 participants by country.

Table 5
Cumulative ordinal model. Individual characteristics and exposition to feedback as predictors of appraisal of teacher feedback.

Predictor		New Zealand						Spain						United States					
		β	p		OR	OR 95 % CI [LL, UL]	Effect Size	β	p		OR	OR 95 % CI [LL, UL]	Effect Size	β	p		OR	OR 95 % CI [LL, UL]	Effect Size
Helpfulness	Type Message ^a	1.737	.000	***	5.683	[5.68, 2.75]	3.133	0.373	.203		1.452	[1.45, 0.95]	0.800	1.395	.000	***	4.036	[4.04, 2.11]	2.225
	Agreeableness	0.334	.196		1.397	[1.4, 0.96]	0.770	0.036	.899		1.037	[1.04, 0.81]	0.571	0.736	.002	**	2.088	[2.09, 1.39]	1.151
	Global Feedback	0.340	.091		1.404	[1.4, 1.03]	0.774	0.492	.000	***	1.635	[1.63, 1.28]	0.901	0.429	.127		1.536	[1.54, 1.01]	0.847
	Feed, Exp (Mod)	−0.892	.046	*	0.410	[0.41, 0.2]	0.226	−0.204	.641		0.815	[0.82, 0.51]	0.449	−0.719	.127		0.487	[0.49, 0.24]	0.269
	Pseudo R ²																		
	McFadden	0.134						0.046						0.155					
	Cox and Snell (ML)	0.282						0.082						0.303					
Usefulness	Type Message ^a	3.343	.000	***	28.295	[28.29, 11.62]	15.600	2.064	.000	***	7.879	[7.88, 4.95]	4.344	2.554	.000	***	12.858	[12.86, 6.41]	7.089
	Conscientiousness	−0.575	.010	*	0.563	[0.56, 0.38]	0.310	−0.073	.755		0.930	[0.93, 0.72]	0.513	0.357	.154		1.429	[1.43, 0.99]	0.788
	Global Feedback	0.057	.816		1.059	[1.06, 0.77]	0.584	0.523	.000	***	1.687	[1.69, 1.32]	0.930	−0.058	.885		0.943	[0.94, 0.64]	0.520
	Pseudo R ²																		
	McFadden	0.227						0.127						0.163					
	Cox and Snell (ML)	0.484						0.270						0.380					
	Intention																		
Intention	Type Message ^a	2.550	.000	***	12.804	[12.8, 6.1]	7.059	1.091	.000	***	2.978	[2.98, 1.91]	1.642	1.949	.000	***	7.018	[7.02, 3.65]	3.869
	Global Feedback	0.316	.141		1.371	[1.37, 0.99]	0.756	0.437	.002	**	1.548	[1.55, 1.21]	0.854	0.115	.712		1.122	[1.12, 0.76]	0.618
	Pseudo R ²																		
	McFadden	0.178						0.065						0.142					
	Cox and Snell (ML)	0.393						0.142						0.311					

Note: * $p < .05$; ** $p < .01$; *** $p < .001$. Values in square brackets indicate the 95% confidence interval of odd ratio estimation. LL and UL indicate the lower and upper limits of the odd ratio. All coefficients are standardized. Coefficients for thresholds and non-significant effects were omitted to streamline the presentation.

Pseudo-R² indices are variations of the R² concept in the OLS regression model; however, none correspond to predicted efficiency or can be tested in the inferential framework (Peng et al., 2002).

^a Type Message: Suggestive/Neutral corresponds to the reference group. All p-values were adjusted the Benjamini–Hochberg correction

Table 6
Cumulative ordinal model. Individual characteristics and exposition to feedback as predictors of appraisal of peer feedback.

Appraisal	Predictor	New Zealand			Spain			United States				
		β	<i>p</i>	OR	OR 95 % CI [LL, UL]	Effect Size	β	<i>p</i>	OR	OR 95 % CI [LL, UL]	Effect Size	
Helpfulness	Type Message ^a	2.162	.000	8.690	[4.31, 17.5]	4.791	0.511	.038	*	1.667	[1.12, 2.49]	0.919
	Global Feedback	0.375	.082	1.456	[1.05, 2.02]	0.802	0.475	.000	***	1.607	[1.27, 2.03]	0.886
	Pseudo R ²											
	McFadden	0.117					0.042				0.131	.131
Usefulness	Cox and Snell	0.287					0.098				0.295	.295
	(ML)											
	Type Message ^a	3.188	.000	24.23	[10.9, 53.7]	13.36	1.692	.000	***	5.430	[3.54, 8.33]	2.994
	Global Feedback	0.380	.063	1.463	[1.07, 2.01]	0.807	0.315	.020	*	1.370	[1.09, 1.72]	0.755
Intention	Pseudo R ²											
	McFadden	0.201					0.085				0.141	
	Cox and Snell	0.455					0.211				0.330	
	(ML)											
	Type Message ^a	2.201	.000	9.032	[4.44, 18.4]	4.979	0.790	.000	***	2.204	[1.48, 3.28]	1.215
	Agreeableness	−0.128	.729	0.880	[0.62, 1.24]	0.485	0.356	.008	**	1.428	[1.14, 1.8]	0.787
	Global Feedback	0.525	.010	1.690	[1.2, 2.37]	0.932	0.151	.386		1.163	[0.93, 1.46]	0.641
	Pseudo R ²											
	McFadden	0.135					0.039				0.147	
	Cox and Snell	0.330					0.105				0.327	
	(ML)											

Note: * $p < .05$; ** $p < .01$; *** $p < .001$. Values in square brackets indicate the 95 % confidence interval of odd ratio estimation. LL and UL indicate the lower and upper limits of the odd ratio. All coefficients are standardized. Coefficients for thresholds and non-significant effects were omitted to streamline the presentation.

Pseudo-R² indices are variations of the R² concept in the OLS regression model; however, none correspond to predicted efficiency or can be tested in the inferential framework (Peng et al., 2002).

^a Type Message: Suggestive/Neutral corresponds to the reference group. All p-values were adjusted the Benjamini-Hochberg correction

those receiving evaluative/positive peer comments. Similar patterns were observed in Spain and New Zealand.

Concerning personality traits, the findings revealed that agreeableness significantly and positively influenced the intent to use peer comments in Spain ($\beta = 0.356$, $p = .008$). Specifically, each standard deviation unit increase in agreeableness led to a 41.8 % increase in the likelihood of intending to use peer comments to enhance the task. Moreover, personality traits did not exhibit significant effects on peer feedback appraisal items in the United States and New Zealand. Furthermore, the results showed a significant and positive effect of global receptivity on the appraisal of peer comments across all countries. In the United States, for every standard deviation unit increase in global feedback, the odds of agreeing that peer messages were helpful and useful increased by 2.58 ($\beta = 0.947$, $p < .001$) and 1.97 ($\beta = 0.676$, $p < .001$) times, respectively, while the likelihood of intending to use peer comments increased by 2.76 ($\beta = 1.014$, $p < .001$) times. In Spain, global feedback had no significant effect on the intent to use peer comments ($\beta = 0.151$, $p > .356$).

Conversely, in New Zealand, for each increase in one standard deviation in global feedback, the probability of increasing the intent to use peer feedback increases by 69% ($\beta = 0.525$, $p = .010$). Similar to results regarding teacher comments, there were no significant effects of age, sex at birth, and GPA on the probability of responses to feedback appraisal due to peer messages. Tables S8 and S9 (Supplementary Materials) present the models that explain the effect of individual variables on emotions elicited by teachers and peer messages, respectively.

3.6. Individual characteristics as predictors of students' preferences for teacher feedback

Regarding the overall perceived value of the message, 72.7% of participants ranked the teacher's message as the most valuable for enhancing their final presentation. This proportion remained consistent across countries, with 70.4% in New Zealand, 75.2% in Spain, and 69.5% in the United States.

However, there were differences in the proportions based on the conditions. In condition 1 (evaluative/positive teacher vs. suggestive/neutral peer), 49.5% of participants rated the teacher's message as the most valuable. This percentage varied across countries, with 53.4% in Spain, 46.4% in the United States, and 39.4% in New Zealand. Conversely, for condition 2 (suggestive/neutral teacher vs. evaluative/positive peer), 94% of the total sample considered the teacher's message as the most valuable. This high proportion was consistent across countries, with 97.5% in New Zealand, 94.6% in Spain, and 93.8% in the United States.

Logistic regression was performed to examine the effect of the experimental condition, country, global receptivity to feedback, and personality on the probability of ranking the teacher comments as the most valuable feedback (Table S9, Supplementary Material). The Hosmer-Lemeshow (H-L) test examined goodness-of-fit inferences. The H-L test ($\chi^2(8) = 5.339$, $p = .720$) suggested that the model fit the data well. The results revealed that, although the country did not have a significant effect on the probability of selecting teacher comments as the most valuable feedback, participants exposed to a suggestive/neutral teacher and evaluative/positive peer feedback were 21 times more likely to rank teacher comments as the most valuable message compared to those exposed to an evaluative/positive teacher and suggestive/neutral peer feedback ($OR = 21.38$ CI[12.2, 36.4], $p < .001$). There were no significant effects on the likelihood of preferring teacher messages based on personal characteristics, personality traits, and feedback beliefs.

4. Discussion

In this experimental vignette study, we investigated students' responses to *discrepant* feedback from a peer and a teacher. We found differences in students' perceptions of utility, helpfulness, and

intentions to use discrepant peer- and teacher-provided feedback. Specifically, a clear preference for teacher feedback over peer feedback emerged from the data, irrespective of the type of message that students were asked to evaluate. However, this evaluation was impacted by the content of the feedback, as 49.5% of participants in the teacher evaluative/positive condition rated the teacher's message as the most valuable, whereas this rate was of 94% when the teacher was the one providing the suggestive/neutral feedback. Interestingly, we observed that students' preferences for teacher feedback were not always in line with their assessments of utility of the message. Although students were generally able to recognize the suggestive and neutral feedback as being more useful than praise only, teacher feedback garnered higher scores irrespective of its type and quality. This underscores the existing discrepancy between perceived usefulness and actual preference, with students valuing feedback from instructors independently of its content.

Earlier studies showed that feedback that provided guidance for improvement was considered more useful than any kind of evaluative comments (Hattie et al., 2021; Weaver, 2006). This relative pattern is most certainly observed in our study, with participants assigning more value for suggestive feedback, compared to evaluative, for both teacher and peer conditions. Their overall *preference* for teacher feedback was notably higher, irrespective of the content. These findings align with existing literature indicating that students typically favor instructor feedback (Van Der Kleij & Lipnevich, 2021). To our knowledge, however, ours was the first study to consider students' relative preferences of peer and teacher feedback of varying content. This finding is troubling because *meta*-analyses have reported significant positive effects of peer feedback on students' performance (Graham et al., 2015; Huisman et al., 2019; Lv et al., 2021). After all, peers represent a tremendous resource in providing more frequent feedback with a high level of comprehensibility, as peer feedback may better align with peers' language proficiency and developmental readiness for understanding the targeted message. For instance, Zhao (2010) found that despite students' higher uptake of teacher feedback, only 58% of teacher feedback instances were applied with a genuine grasp of their need and importance, whereas that rate was 83% for peer feedback. Additionally, peer feedback can save educators time that could be devoted to other instructional activities. However, although students recognize the value of feedback, there remains a prevalent tendency to prioritize teacher feedback.

Our second research question investigated whether the *discrepant* peer- and teacher-provided feedback would elicit eight different emotions in students. Discrete emotions were divided into positive (happiness, pride, enthusiasm, inspiration) and negative (distress, fear, irritation, and feeling upset) (Watson et al., 1988). Interestingly, despite the higher ratings of helpfulness and usefulness that students gave to the neutral suggestive feedback, the positive evaluative message elicited the highest ratings of positive emotions. This finding further supports the claim that positive feedback that indicates success can enhance perceived control, thus leading to an increase in positive emotional experiences like enthusiasm and pride.

A main effect of source on positive emotions was also observed, with praise from the teacher leading to higher positive emotions than receiving the same positive feedback from a peer. Therefore, in addition to the preference for teacher feedback, we also observed teachers' critical role in eliciting enhanced positive emotional experiences in students. It is noteworthy that despite the current research interest in various forms of feedback (i.e., computer, peers, rubric) and the positive impact all those sources have demonstrated on student achievement (Lipnevich et al., 2023; Tomazin et al., 2023), they do not replace the benefits of teacher comments for the teacher-student relationship, as teacher feedback is an essential trigger for positive academic emotions (Hill et al., 2021).

Moreover, we also found that participants who received praise were more likely to report intense feelings for all positive emotions than those who received a neutral-suggestive message. Unfortunately, despite the

positive effect of positive feedback on students' sense of worth and confidence (Hill et al., 2021), this form of feedback does not always enhance motivation or performance (Hill et al., 2021; Lipnevich & Smith, 2009a).

Regarding experiences of negative emotions, students reported higher rates of negative emotions in the suggestive/neutral condition compared to the evaluative/positive one. Furthermore, peer feedback in the suggestive/neutral also triggered higher ratings of negative emotions than teacher feedback in the same message condition. In fact, 28.7% of participants expressed feeling distressed after receiving feedback from peers. Those results point to a higher tolerance for receiving suggestions from teachers, as those are seen as the authority, and students expect to be corrected by them, compared to peers, who may be viewed as less credible/trusted sources. A significant body of literature has emphasized the importance of the source's credibility for students' engagement with the feedback (Alqassab et al., 2023; Poulos & Mahony, 2008; Tsui & Ng, 2000). For instance, Kaufman & Schunn (2011) found that students' distrust in the fairness and quality of feedback triggered emotions such as distrust and doubt.

Regarding appraisals of teacher feedback, we found that personality traits influenced these evaluations differently across countries after controlling by other individual and academic characteristics (age, gender, and GPA). In the United States, agreeableness positively affected the perceived helpfulness of teacher feedback, while in New Zealand, conscientiousness negatively affected the usefulness of teacher feedback. Further, global receptivity to feedback positively affects feedback appraisals in Spain, increasing the likelihood of perceiving teacher feedback as helpful, useful, and likely to be used. However, moderate exposure to teacher feedback reduced its perceived helpfulness in New Zealand compared to lower exposure levels.

In the context of peer feedback evaluations, agreeableness was positively associated with the intention to utilize peer feedback in Spain, whereas personality traits did not exhibit significant effects in the United States or New Zealand. Moreover, global receptivity contributed to more favorable appraisals of peer feedback, increasing the likelihood of perceiving it as beneficial, relevant, and applicable. However, in Spain, global receptivity did not influence the intention to incorporate peer feedback. That is, individuals who valued more receiving feedback tended to rate the comments provided by different sources (i.e., teacher and peers) as higher in utility and helpfulness. High receptivity to feedback was also linked to higher reports of positive emotions and lower rates of negative emotions. Those results are in accordance with Lipnevich and colleagues (Lipnevich, Gjicali, Asil, & Smith, 2021; Lipnevich & Lopera-Oquendo, 2024).

Our findings also revealed that greater exposure to feedback seems to bolster appraisals, implying that familiarity with feedback mechanisms may influence how individuals interpret and respond to feedback. This is a major finding and contribution from this study, as it demonstrates the importance of increasing feedback practices in classes so students may better understand the importance of feedback to enhance their performance. It is also in accordance with Wu and Schunn (2020)'s finding that the frequency at which feedback is given significantly impacts how much students make use of peer feedback, as repeated references to feedback result in higher rates of implementation.

Finally, our fourth research question examined whether our findings could be generalized among the three participating countries of Spain, New Zealand, and the United States. The general trend remained consistent across the three countries. In other words, students preferred teacher provided feedback irrespective of its content. Positive emotions were also stronger for teacher feedback compared to peer feedback, and negative emotions were higher for peer feedback. However, we observed some noteworthy differences, particularly in the low ratings of peer feedback among New Zealand students. As revealed in our study, the experience of receiving feedback plays a significant role in shaping individuals' appraisals of that feedback. It is possible that participants in our sample have had little opportunity to engage in peer feedback

practice, as these results are contrary to the findings shown by [Barnard et al. \(2015\)](#), which suggested that despite their New Zealand participants' initial concern about engaging in peer feedback practices, they reported peer feedback to be useful. Further research is needed to determine the length of exposure needed to have a meaningful impact on feedback appraisal. Overall, these findings highlight different dynamics involved in how feedback messages and their sources impact the students' responses and preferences for feedback, underscoring the importance of considering contextual and individual factors in understanding the complexity behind feedback's impact on learning in educational settings.

4.1. Practical implications

The findings from this study suggest a range of important practical implications for professional development and for communicating feedback to students. Specifically, there appears to be a disconnect between students' preferences for evaluative, positively-valence feedback and the limited empirical evidence on praise's effectiveness for promoting actual improvement and learning (e.g., [Lipnevich et al., 2023](#)). Although students in our sample reported experiencing high positive emotions in response to receiving praise, decades of research have demonstrated that praise can be an ineffective feedback approach for fostering substantive growth ([Lipnevich, Eßer, Park, & Winstone, 2023](#); [Lipnevich & Smith, 2009b](#)).

Praise statements lack the important information students need to identify areas for improvement and adjust their strategies accordingly ([Hattie & Zierer, 2019](#)). As such, a key recommendation for educators is to receive training on providing feedback that prioritizes specificity, clarity, and actionable guidance over positively framed comments. Professional development should equip instructors with strategies for objectively identifying skill strengths and challenges, as well as providing explicit suggestions for concrete next steps to make performance gains while communicating feedback in a manner that balances candor with sensitivity. Additionally, educating students themselves on the limitations of praise and benefits of suggestive actionable comments could be valuable. Building students' receptivity to feedback could enhance their ability to view feedback as an opportunity versus a personal criticism.

Another key practical implication relates to students' tendency to regard peer feedback as less useful than instructor feedback, irrespective of the quality of peer comments. Instructors should make focused efforts to emphasize the pedagogical value of high-quality peer feedback. Although students may initially dismiss it due to perceived authority or credibility issues, peer feedback holds significant potential benefits by exposing students to diverse perspectives beyond just the instructor's lens. To unlock these learning opportunities, instructors should directly address common skepticisms about peer feedback through class discussions. They can provide training on strategies for both giving and receiving high-quality peer feedback and modeling constructive examples. Creating accountability by incorporating peer feedback into grading can further signal its importance. Facilitating reflections where students analyze the strengths of received peer feedback could also build appreciation for this rich resource. By proactively advocating for peer feedback's merits, instructors can reshape student view to be more receptive to this valuable channel for improvement.

To this end, our findings revealed that exposure to peer and teacher feedback was an important predictor of positive attitudes toward peer feedback. Hence, a third practical implication is that instructors should provide ample opportunities for students to experience and engage with peer feedback processes. Increased exposure and first-hand experience seem critical for combating negative preconceptions about the value of peer input. By intentionally incorporating peer feedback activities and assignments throughout a course, instructors can foster greater acceptance and appreciation for this feedback source over time. Encouraging reciprocal peer review, where students both give and receive feedback

from peers, could be particularly powerful. As students practice evaluating others' work while receiving feedback themselves, they may develop more calibrated perceptions of peer feedback quality.

4.2. Limitations and future directions

One limitation of the present study is related to the distinction between feedback preference and behavior. In this study, we gauged student emotions, appraisals, and intention to use feedback through vignettes rather than measuring feedback-related behavior directly. Although several studies have shown strong relationships between intention and behavior, other factors such as school culture, teacher-peer relationships, and peer-to-peer connections may moderate the relationship between intention and behavior. Moreover, using vignettes in experimental conditions offers a viable and well-established method for examining complex phenomena, such as feedback perception, under controlled conditions. It has been a widely accepted and validated approach in the field of work psychology and personnel selection ([Oostrom, De Soete, & Lievens, 2015](#); [Webster, Paton, Crampton, & Tiffin, 2020](#)) and recently in education (e.g., [Yerdelen, Durksen, & Klassen, 2018](#); [Murano et al., 2021](#)). Vignettes allow us to isolate specific variables—in our case, feedback content and tone—while minimizing the influence of extraneous factors that are difficult to control in real-world settings. This method provides valuable advantages for evaluating non-cognitive traits, thanks to their high face and predictive validity ([Gold & Holodyski, 2015](#); [Lievens & Sackett, 2012](#); [Mussel, Gatzka, & Hewig, 2016](#)). Moreover, to enhance the realism of our vignettes, we incorporated specific details and cues reflective of common educational settings, ensuring ecological validity to the greatest extent possible (as per recommendations of situational judgment test (SJT) development; see, for example, [Lievens and Sackett \(2012\)](#)). Thus, we contend that vignettes remain a robust and practical tool for investigating the dynamics of feedback interactions in a controlled yet meaningful manner. In summary, using vignettes or SJT as an assessment tool is also less prone to personal biases, as the most socially desirable response is not always apparent ([Lipnevich, MacCann, & Roberts, 2013](#)). Therefore, this approach should not hinder the generalizability of our results to natural settings. This is evidenced by the alignment of our main findings with previous research, suggesting that our controlled approach has not compromised the validity of our conclusions. However, future studies can replicate our research in a more realistic academic situation, for example using an actual task and measuring student behaviors and preferences related to it.

Using single-item measures as outcomes could also be a limitation in our study. While single-item measures in emotions research have been criticized for their drawbacks, such as lower reliability and limited capacity for providing a proper construct coverage ([Allen et al., 2022](#)), this approach is common in emotions research. It has shown satisfactory reliability and correlation with original emotion scales ([Bieleke et al., 2021](#)). In our study, the use of single-item measures was also supported by the lack of measurement equivalence across countries for Positive Affect (PA) and Negative Affect (NA) latent variables, which has also been supported by the empirical evidence provided by [Wedderhoff et al. \(2021\)](#). Regarding feedback appraisal single items approach, there is insufficient theoretical support to assume that these items are part of a unique construct.

Consequently, while latent variable modeling is valuable for understanding underlying structures, our findings and research goals support using single-item measures for emotions and feedback appraisal as a valid and practical approach in this study without compromising the predictive validity of emotion measures. Furthermore, using single-item measures reduces the risk of introducing cultural biases that may arise from differential item functioning ([Abdel-Khalek, 2006](#); [Dejonckheere et al., 2022](#)). Future studies should consider adopting comprehensive, validated latent measures that tap into the different families of emotional experiences to map how feedback interacts with students'

affect more precisely. Complementing self-reports with classroom observations or physiological indices could further strengthen construct measurement.

A potential concern relates to constructing our feedback messages and its implications for result interpretation. While the simultaneous variation in feedback content (evaluative vs. suggestive) and tone (positive vs. neutral) may complicate isolating the influence of each factor on students' evaluations, our study design intentionally reflects the complexity of real-world feedback. In practice, feedback rarely varies along a single dimension; instead, content and tone interact to shape recipients' perceptions and responses (see [Máñez, Lipnevich, Lopera-Oquendo, & Cerdán, 2024](#)). Our objective was to examine how these naturally occurring combinations influence students' reactions rather than artificially isolating each factor, which would be impractical (feedback inherently contains content). Regarding potential language differences between peer and teacher feedback, we acknowledge concerns about confounding factors. However, linguistic variations were minimal. The messages maintained near-identical content and structure, with only slight adjustments to reflect natural differences in communication style between peers and teachers. These subtle modifications were intentionally included to align with authentic feedback dynamics in educational settings, ensuring consistency in core messaging while preserving ecological validity. In sum, by incorporating these minimal adjustments, our study better reflects the natural dynamics of feedback exchanges, enhancing ecological validity. Another limitation (or rather a warning for careful interpretation of findings) concerns the composition of our samples. Although the participants were drawn from higher education settings, the three samples exhibit notable differences in age and gender, which complicates direct comparisons. Nevertheless, the inclusion of diverse countries in the study is well-justified, and we have tried to strengthen the rationale behind this choice. Spain, New Zealand, and the United States each represent distinct educational, cultural, and policy contexts, which enhances the value of their inclusion in cross-national research examining global educational trends. Spain offers insights into systems shaped by European Union policies and strongly emphasizes regional autonomy, including linguistic and curricular variations. New Zealand provides a unique perspective as a small Pacific nation operating within a bicultural framework established by the Treaty of Waitangi, focusing on equity for Māori and Pasifika students and a competency-based national curriculum. The United States, with its vast size and decentralized educational system, presents a blend of federal oversight and local control, marked by significant socioeconomic and cultural diversity. As such, including these three countries facilitates a meaningful comparative analysis ([Marín et al., 2022](#); [Tan & Pillay, 2008](#)). Future research could replicate this study using matched groups of students to further explore the generalizability of these findings, including both the similarities and differences observed across countries.

4.3. Conclusion

The aim of this experimental study was to explore the potential influence of discrepant feedback between teachers and peers on emotional responses and appraisal to feedback messages. Utilizing a cross-country sample of post-secondary students encompassing Spain, the United States, and New Zealand, we explored individual and contextual factors, such as personality, receptivity to feedback, and exposure to teacher and peer feedback, that could explain variations in how students perceive and react to feedback within academic settings.

Our findings underscore a clear preference among students for teacher feedback over peer feedback, irrespective of the message focus and valence (evaluative/neutral vs. suggestive/positive). Despite recognizing the utility of suggestive/neutral feedback, students consistently valued feedback from instructors more and this pattern persisted across the three countries. This aligns with existing literature emphasizing the importance students place on teacher feedback for guidance

and improvement. Furthermore, our study explored how different emotional experiences were triggered by discrepant feedback. Positive evaluative messages elicited higher positive emotions, reinforcing the role of positive feedback in enhancing perceived control and fostering enthusiasm and pride. Interestingly, despite the benefits of peer feedback, students reported more positive emotions and tolerance for suggestive messages when this type of message is received from teachers.

In addition to feedback source and message valence, our analysis revealed that personality traits, receptivity to feedback, and exposure to feedback played significant roles in students' responses to feedback. Practical implications include the need for targeted training on feedback provision, emphasizing the pedagogical value of peer feedback, and creating opportunities for students to engage with diverse feedback sources. Overall, our study is among the first ones to consider how students may negotiate discrepant feedback messages from multiple sources. We hope that it will inspire future studies and offer insights for improving feedback practices in educational settings.

CRedit authorship contribution statement

Carolina Lopera-Oquendo: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Anastasiya A. Lipnevich:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Ligia Tomazin:** Writing – review & editing, Writing – original draft, Conceptualization. **Ignacio Máñez:** Writing – review & editing. **Samuel P. León:** Writing – review & editing. **Nicola Beatson:** Writing – review & editing, Data curation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cedpsych.2025.102394>.

References

- Abdel-Khalek, A. (2006). Measuring happiness with a single-item scale. *Social Behavior and Personality*, 34, 139–150. <https://doi.org/10.2224/SBP.2006.34.2.139>
- Aben, J. E. J., Timmermans, A. C., Dingyloudi, F., Lara, M. M., & Strijbos, J.-W. (2022). What influences students' peer-feedback uptake? relations between error tolerance, feedback tolerance, writing self-efficacy, perceived language skills and peer-feedback processing. *Learning and Individual Differences*, 97, Article 102175. <https://doi.org/10.1016/j.lindif.2022.102175>
- Adams, A.-M., Wilson, H., Money, J., Palmer-Conn, S., & Fearn, J. (2020). Student engagement with feedback and attainment: The role of academic self-efficacy. *Assessment & Evaluation in Higher Education*, 45(2), 317–329. <https://doi.org/10.1080/02602938.2019.1640184>
- Agius, N. M., & Wilkinson, A. (2014). Students' and teachers' views of written feedback at undergraduate level: A literature review. *Nurse Education Today*, 34(4), 552–559. <https://doi.org/10.1016/j.nedt.2013.07.005>
- Albright, M. D., & Levy, P. E. (1995). The effects of source credibility and performance rating discrepancy on reactions to multiple raters. *Journal of Applied Social Psychology*, 25(7), 577–600. <https://doi.org/10.1111/j.1559-1816.1995.tb01600.x>
- Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science: A call to action [editorial]. *European Journal of Psychological Assessment*, 38(1), 1–5. <https://doi.org/10.1027/1015-5759/a000699>
- Alqassab, M., Strijbos, J.-W., Panadero, E., Ruiz, J. F., Warrens, M., & To, J. (2023). A systematic review of peer assessment design elements. *Educational Psychology Review*, 35(1), 18. <https://doi.org/10.1007/s10648-023-09723-7>
- Baadte, C., & Schnotz, W. (2014). Feedback effects on performance, motivation and mood: Are they moderated by the learner's self-concept? *Scandinavian Journal of Educational Research*, 58(5), 570–591. <https://doi.org/10.1080/00313831.2013.781059>
- Barnard, R., de Luca, R., & Li, J. (2015). First-year undergraduate students' perceptions of lecturer and peer feedback: A New Zealand action research project. *Studies in Higher Education*, 40(5), 933–944. <https://doi.org/10.1080/03075079.2014.881343>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate—A Practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B—Methodological*, 57, 289–300.
- Bieleke, M., Gogol, K., Goetz, T., Daniels, L., & Pekrun, R. (2021). The AEQ-S: A short version of the achievement emotions questionnaire. *Contemporary Educational Psychology*, 65, Article 101940. <https://doi.org/10.1016/j.cedpsych.2020.101940>
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2), 230–258. <https://doi.org/10.1177/0049124192021002005>

- Brown, G. T. L., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28(7), 968–978. <https://doi.org/10.1016/j.tate.2012.05.003>
- Brown, G., Harris, L., & Harnett, J. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*. <https://doi.org/10.1007/s11092-013-9187-5>
- Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science*, 2(1), 77–101. <https://doi.org/10.1177/2515245918823199>
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105(3), 456–466. <https://doi.org/10.1037/0033-2909.105.3.456>
- Chen, F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Cheng, K.-H., Liang, J.-C., & Tsai, C.-C. (2015). Examining the role of feedback messages in undergraduate students' writing performance during an online peer assessment activity. *The Internet and Higher Education*, 25, 78–84. <https://doi.org/10.1016/j.iheduc.2015.02.001>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. https://doi.org/10.1207/s15328007SEM0902_5
- Dejonckheere, E., Demeyer, F., Geusens, B., Piot, M., Tuerlinckx, F., Verdonck, S., & Mestdagh, M. (2022). Assessing the reliability of single-item momentary affective measurements in experience sampling. *Psychological assessment*. <https://doi.org/10.1037/pas0001178>
- Dijks, M. A., Brummer, L., & Kostons, D. (2018). The anonymous reviewer: the relationship between perceived expertise and the perceptions of peer feedback in higher education. *Assessment & Evaluation in Higher Education*, 43(8), 1258–1271. <https://doi.org/10.1080/02602938.2018.1447645>
- Ekoniak, M., & Paretti, M. C. (2018). Instructor vs peer writing feedback in a large first-year engineering course. In *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1–8). <https://doi.org/10.1109/FIE.2018.8659050>
- Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment? *BMC medical research methodology*, 2, 8. <https://doi.org/10.1186/1471-2288-2-8>
- Floress, M. T., Beschta, S. L., Meyer, K. L., & Reinke, W. M. (2017). Praise research trends and future directions: Characteristics and teacher training. *Behavioral Disorders*, 43(1), 227–243. <https://doi.org/10.1177/0198742917704648>
- French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling*, 13(3), 378–402. https://doi.org/10.1207/s15328007sem1303_3
- Gallardo-Pujol, D., Rouco, V., Cortijos-Bernabeu, A., Ocejá, L., Soto, C. J., & John, O. P. (2022). Factor structure, general invariance, measurement properties, and short forms of the Spanish adaptation of the Big Five Inventory-2. *Psychological Test. Adaptation and Development*, 3(1), 44–69. <https://doi.org/10.1027/2698-1866/a000020>
- Gan, Z., Liu, F., & Nang, H. (2023). The role of self-efficacy, task value, and intrinsic and extrinsic motivations in students' feedback engagement in english learning. *Behavioral Sciences*, 13(5). <https://doi.org/10.3390/bs13050428>. Article 5.
- Gold, B., & Holodynski, M. (2015). Development and construct validation of a situational judgment test of strategic knowledge of classroom management in elementary schools. *Educational Assessment*, 20(3), 226–248. <https://doi.org/10.1080/10627197.2015.1062087>
- Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, 115(4), 523–547. <https://doi.org/10.1086/681947>
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, 44(11 Suppl 3), S78–S94. <https://doi.org/10.1097/01.mlr.0000245454.12228.8f>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.
- Harris, L. R., Brown, G. T. L., & Harnett, J. A. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*, 26(2), 107–133. <https://doi.org/10.1007/s11092-013-9187-5>
- Harris, L. R., Brown, G. T. L., & Dargusch, J. (2018). Not playing the game: Student assessment resistance as a form of agency. *The Australian Educational Researcher*, 45(1), 125–140. <https://doi.org/10.1007/s13384-018-0264-0>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 88–118. <https://doi.org/10.3102/003465430298487>
- Hattie, J., & Zierer, K. (2019). *Visible learning insights*. Routledge. <https://doi.org/10.4324/9781351002226>
- Hattie, J., Crivelli, J., Van Gompel, K., West-Smith, P., & Wike, K. (2021). Feedback that leads to improvement in student essays: Testing the hypothesis that “where to next” feedback is most powerful. *Frontiers in Education*, 6. <https://www.frontiersin.org/articles/10.3389/feduc.2021.645758>
- Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, 128(5), 774–795. <https://doi.org/10.1037/0033-2909.128.5.774>
- Hill, J., Berlin, K., Choate, J., Cravens-Brown, L., McKendrick-Calder, L., & Smith, S. (2021). Exploring the emotional responses of undergraduate students to assessment feedback: Implications for instructors. *Teaching & Learning Inquiry*, 9(1), 294–316. <https://doi.org/10.20343/teachlearninqu.9.1.20>
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037/1082-989X.3.4.424>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Huisman, B., Saab, N., Van Den Broek, P., & Van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education*, 44(6), 863–880. <https://doi.org/10.1080/02602938.2018.1545896>
- Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, 14(1), 63–76. <https://doi.org/10.1177/1469787412467125>
- Jonsson, A., & Panadero, E. (2018). Facilitating students' active engagement with feedback. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (1st ed, pp. 531–553). Cambridge University Press. <https://doi.org/10.1017/9781316832134.026>
- Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science*, 39(3), 387–406. <https://doi.org/10.1007/s11251-010-9133-6>
- Kerman, N. T., Noroozi, O., Banihashem, S. K., Karami, M., & Biemans, H. J. A. (2022). Online peer feedback patterns of success and failure in argumentative essay writing. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2022.2093914>
- Lievens, F., & Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, 97(2), 460–468. <https://doi.org/10.1037/a0025741>
- Lipnevich, A. A., & Lopera-Oquendo, C. (2024). Receptivity to instructional feedback: A validation study in the secondary school context in Singapore. *European Journal of Psychological Assessment*, 40(1), 22–32. <https://doi.org/10.1027/1015-5759/a000733>
- Lipnevich, A. A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *Frontiers in Education*, 6, Article 720195. <https://doi.org/10.3389/feduc.2021.720195>
- Lipnevich, A. A., & Smith, J. K. (2009a). Effects of differential feedback on students' examination performance. *Journal of Experimental Psychology: Applied*, 15(4), 319–333. <https://doi.org/10.1037/a0017841>
- Lipnevich, A. A., & Smith, J. K. (2009b). “I really need feedback to learn”: Students' perspectives on the effectiveness of the differential feedback messages. *Educational Assessment, Evaluation and Accountability*, 21(4), 347–367. <https://doi.org/10.1007/s11092-009-9082-2>
- Lipnevich, A. A., & Smith, J. K. (2022). Student – feedback interaction model: Revised. *Studies in Educational Evaluation*, 75, Article 101208. <https://doi.org/10.1016/j.stueduc.2022.101208>
- Lipnevich, A. A., MacCann, C., & Roberts, R. D. (2013). Assessing noncognitive constructs in education: A review of traditional and innovative approaches. In D. H. Saklofske, C. B. Reynolds, & V. L. Schwane (Eds.), *Oxford handbook of child psychological assessment* (pp. 750–772). Cambridge, MA: Oxford University Press.
- Lipnevich, A. A., Gjicali, K., Asil, M., & Smith, J. K. (2021). Development of a measure of receptivity to instructional feedback and examination of its links to personality. *Personality and Individual Differences*, 169, Article 110086. <https://doi.org/10.1016/j.paid.2020.110086>
- Lipnevich, A. A., Murano, D., Krannich, M., & Goetz, T. (2021). Should I grade or should I comment: Links among feedback, emotions, and performance. *Learning and Individual Differences*, 89, Article 102020. <https://doi.org/10.1016/j.lindif.2021.102020>
- Lipnevich, A. A., Eßer, F. J., Park, M. J., & Winstone, N. (2023). Anchored in praise? potential manifestation of the anchoring bias in feedback reception. *Assessment in Education: Principles, Policy & Practice*, 30(1), 4–17. <https://doi.org/10.1080/0969594X.2023.2179956>
- Lugtj, P., Boeije, H. R., & Lensvelt-Mulders, G. J. L. M. (2012). Change? what change? *Methodology*, 8(3), 115–123. <https://doi.org/10.1027/1614-2241/a000043>
- Lv, X., Ren, W., & Xie, Y. (2021). The effects of online feedback on ESL/EFL writing: A meta-analysis. *The Asia-Pacific Education Researcher*, 30(6), 643–653. <https://doi.org/10.1007/s40299-021-00594-6>
- MacCallum, R. C., Browne, M. W., Sugawara, H. M., & Wegener, D. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130–149. <https://doi.org/10.1037/1082-989X.1.2.130>
- Mahvelati, E. (2021). Learners' perceptions and performance under peer versus teacher corrective feedback conditions. *Studies in Educational Evaluation*, 70. <https://doi.org/10.1016/j.stueduc.2021.100995>
- Mandouit, L., & Hattie, J. (2023). Revisiting “the power of feedback” from the perspective of the learner. *Learning and Instruction*, 84, Article 101718. <https://doi.org/10.1016/j.learninstruc.2022.101718>
- Máñez, I., Lipnevich, A. A., Lopera-Oquendo, C., & Cerdán, R. (2024). Examining pre-service teachers' feedback on low- and high-quality written assignments. *Educational Assessment, Evaluation and Accountability*, 36(2), 225–256. <https://doi.org/10.1007/s11092-024-09432-x>
- Máñez, I., Skrobiszewska, N., Descals, A., Cantero, M. J., Cerdán, R., García, Ó. F., & García-Ros, R. (2024). Channelling feedback through audiovisual presentations: Do higher education students perceive, use and benefit from video feedback compared to written feedback? *Journal of Computer Assisted Learning*, 40(4), 1886–1897. <https://doi.org/10.1111/jcal.12993>

- Marín, V. I., Zawacki-Richter, O., Aydin, C. H., Bedenlier, S., Bond, M., Bozkurt, A., Conrad, D., Jung, I., Kondakci, Y., Prinsloo, P., Roberts, J., Veletsianos, G., Xiao, J., & Zhang, J. (2022). Faculty perceptions, awareness and use of open educational resources for teaching and learning in higher education: A cross-comparative analysis. *Research and Practice in Technology Enhanced Learning*, 17(11). <https://doi.org/10.1186/s41039-022-00185-z>
- Metcalfe, J., Kornell, N., & Finn, B. (2009). Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & Cognition*, 37(8), 1077–1087. <https://doi.org/10.3758/MC.37.8.1077>
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111–130. <https://doi.org/10.21500/20112084.857>
- Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, 39(3), 479–515. <https://doi.org/10.1207/S15327906MBR3903.4>
- Morris, P., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3), e3292. <https://doi.org/10.1002/rev3.3292>
- Murano, D., Lipnevich, A. A., Walton, K. E., Burrus, J., Way, J. D., & Anguiano-Carrasco, C. (2021). Measuring social and emotional skills in elementary students: Development of self-report Likert, situational judgment test, and forced choice items. *Personality and Individual Differences*, 169, Article 110012. <https://doi.org/10.1016/j.paid.2020.110012>
- Mussel, P., Gatzka, T., & Hewig, J. (2016). Situational judgment tests as an alternative measure for personality assessment. *European Journal of Psychological Assessment*. <https://doi.org/10.1027/1015-5759/a000346>
- Nakata, T. (2015). Effects of feedback timing on second language vocabulary learning: Does delaying feedback increase learning? *Language Teaching Research*, 19(4), 416–434. <https://doi.org/10.1177/1362168814541721>
- Narciss, S., Sosnovsky, S., Schnaubert, L., Andrés, E., Eichmann, A., Goguadze, G., & Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education*, 71, 56–76. <https://doi.org/10.1016/j.compedu.2013.09.011>
- Narciss, S., Prescher, C., Khalifah, L., & Kördle, H. (2022). Providing external feedback and prompting the generation of internal feedback fosters achievement, strategies and motivation in concept learning. *Learning and Instruction*, 82, Article 101658. <https://doi.org/10.1016/j.learninstruc.2022.101658>
- Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment & Evaluation in Higher Education*, 46(5), 756–778. <https://doi.org/10.1080/02602938.2020.1823314>
- OECD. (2019). PISA 2018 Results (Volume I): What students know and can do. OECD. <https://doi.org/10.1787/5f07c754-en>
- OECD. (2023). PISA 2022 Results (Volume I): The state of learning and equity in education. OECD. <https://doi.org/10.1787/53f23881-en>
- Oostrom, J. K., De Soete, B., & Lievens, F. (2015). Situational judgment testing: A review and some new developments. In I. Nikolaou, & J. K. Oostrom (Eds.), *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice* (pp. 172–189). Routledge/Taylor & Francis Group.
- Paterson, C., Paterson, N., Jackson, W., & Work, F. (2020). What are students' needs and preferences for academic feedback in higher education? A systematic review. *Nurse Education Today*, 85, Article 104236. <https://doi.org/10.1016/j.nedt.2019.104236>
- Pekrun, R., Cusack, A., Murayama, K., Elliot, A. J., & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction*, 29, 115–124. <https://doi.org/10.1016/j.learninstruc.2013.09.002>
- Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: The students' perspective. *Assessment & Evaluation in Higher Education*, 33(2), 143–154. <https://doi.org/10.1080/02602930601127869>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Quesada Serra, V., Rodríguez-Gómez, G., & Ibarra-Sáiz, M. (2016). What are we missing? Spanish lecturers' perceptions of their assessment practices. *Innovations in Education and Teaching International*, 53, 48–59. <https://doi.org/10.1080/14703297.2014.930353>
- Ramani, S., Könings, K. D., Ginsburg, S., & van der Vleuten, C. P. M. (2019). Feedback redefined: Principles and practice. *Journal of General Internal Medicine*, 34(5), 744–749. <https://doi.org/10.1007/s11606-019-04874-2>
- R Core Team. (2024). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Rotsaert, T., Panadero, E., Estrada, E., & Schellens, T. (2017). How do students perceive the educational value of peer assessment in relation to its social nature? a survey study in Flanders. *Studies in Educational Evaluation*, 53, 29–40. <https://doi.org/10.1016/j.stueduc.2017.02.003>
- Ruegg, R. (2015). Differences in the uptake of peer and teacher feedback. *RELJ Journal*, 46(2), 131–145. <https://doi.org/10.1177/0033688214562799>
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34(4, Pt. 2), 100.
- Samejima, F. (1997). Graded response model. In W. V. van der Linden, & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). New York, NY: Springer.
- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 167–180. <https://doi.org/10.1080/10705511.2014.882658>
- Schoot, R., Lugtig, P., & Hox, J. (2012). Developmetrics: A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9, 486–492. <https://doi.org/10.1080/17405629.2012.686740>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Sun, H., & Wang, M. (2022). Effects of teacher intervention and type of peer feedback on student writing revision. *Language Teaching Research*. <https://doi.org/10.1177/13621688221080507>
- TALIS (2024). 2018 Results (Volume I). teachers and school leaders as lifelong learners. OECD. Retrieved from March 12. <https://www.oecd.org/education/talis/talis-2018-results-volume-i-1d0bc92a-en.htm>
- Tan, P. L., & Pillay, H. (2008). Understanding learning behavior of Malaysian adult learners: A cross-cultural sensitive framework. *Educational Research for Policy and Practice*, 7, 85–97.
- Thalmayer, A. G., & Saucier, G. (2014). The questionnaire Big Six in 26 nations: Developing cross-culturally applicable Big Six, Big Five and Big Two Inventories. *European Journal of Personality*, 28(5), 482–496. <https://doi.org/10.1002/per.1969>
- Tomazin, L., Lipnevich, A. A., & Lopera-Oquendo, C. (2023). Teacher feedback vs. annotated exemplars: Examining the effects on middle school students' writing performance. *Studies in Educational Evaluation*, 78, Article 101262. <https://doi.org/10.1016/j.stueduc.2023.101262>
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147–170. [https://doi.org/10.1016/S1060-3743\(00\)00022-9](https://doi.org/10.1016/S1060-3743(00)00022-9)
- Van Boekel, M., Hufnagle, A. S., Weisen, S., & Troy, A. (2023). The feedback I want versus the feedback I need: Investigating students' perceptions of feedback. *Psychology in the Schools*, 60(9), 3389–3402. <https://doi.org/10.1002/pits.22928>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–69. <https://doi.org/10.1177/109442810031002>
- Van de Ridder, J. M. M., Berk, F., Stokking, K., & ten Cate, O. (2014). Feedback providers' credibility impacts students' satisfaction with feedback and delayed performance. *Medical Teacher*, 37. <https://doi.org/10.3109/0142159X.2014.970617>
- Van der Kleij, F. M., & Lipnevich, A. A. (2021). Student perceptions of assessment feedback: A critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, 33(2), 345–373. <https://doi.org/10.1007/s11092-020-09331-x>
- Vattøy, K.-D., Gamlem, S. M., & Rogne, W. M. (2021). Examining students' feedback engagement and assessment experiences: A mixed study. *Studies in Higher Education*, 46(11), 2325–2337. <https://doi.org/10.1080/03075079.2020.1723523>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of personality and social psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Weaver, M. R. (2006). Do students value feedback? student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*, 31(3), 379–394. <https://doi.org/10.1080/02602930500353061>
- Webster, E. S., Paton, L. W., Crampton, P. E. S., & Tiffin, P. A. (2020). Situational judgement test validity for selection: A systematic review and meta-analysis. *Medical Education*, 54(10), 888–902. <https://doi.org/10.1111/medu.14201>
- Wedderhoff, N., Gnabts, T., Wedderhoff, O., Burgard, T., & Bošnjak, M. (2021). On the structure of affect: A meta-analytic investigation of the dimensionality and the cross-national applicability of the Positive and Negative Affect Schedule (PANAS). *Zeitschrift für Psychologie*, 229(1), 24–37. <https://doi.org/10.1027/2151-2604/a000434>
- Wingate, U. (2010). The impact of formative feedback on the development of academic writing. *Assessment & Evaluation in Higher Education*, 35(5), 519–533. <https://doi.org/10.1080/02602930903512909>
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of reciprocity processes. *Educational Psychologist*, 52(1), 17–37. <https://doi.org/10.1080/00461520.2016.1207538>
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.03087>
- Wu, Y., & Schunn, C. D. (2020). When peers agree, do students listen? the central role of feedback quality and feedback frequency in determining uptake of feedback. *Contemporary Educational Psychology*, 62, Article 101897. <https://doi.org/10.1016/j.cedpsych.2020.101897>
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179–200. <https://doi.org/10.1016/j.jslw.2006.09.004>
- Yerdelen, S., Durksen, T., & Klassen, R. M. (2018). An international validation of the Engaged Teacher Scale. *Teachers and Teaching: Theory and Practice*, 24(6), 673–689. <https://doi.org/10.1080/13540602.2018.1457024>
- Yu, C. (2002). Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes (Order No. 3066425). Available from ProQuest Dissertations & Theses Global (276287121).

- Zacharias, N. T. (2007). Teacher and student attitudes toward teacher feedback. *REL C Journal*, 38(1), 38–52. <https://doi.org/10.1177/0033688206076157>
- Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a chinese english writing classroom. *Assessing Writing*, 15(1), 3–17. <https://doi.org/10.1016/j.asw.2010.01.002>
- Zumbrunn, S., Marrs, S., & Mewborn, C. (2016). Toward a better understanding of student perceptions of writing feedback: A mixed methods study. *Reading and Writing*, 29(2), 349–370. <https://doi.org/10.1007/s11145-015-9599-3>