



ARTICLE



<https://doi.org/10.1057/s41599-025-04927-4>

OPEN

# Development and validation of a generic self-assessment scale for K-12 teachers as feedback givers: Insights from item response theory and factor analysis

Cui Kang<sup>1,2,3</sup>, Jinyan Huang<sup>1,2,3✉</sup>, Ying Liu<sup>1,2</sup> & Hua Yin<sup>1</sup>

This study aimed to develop and validate a generic self-assessment scale for Chinese K-12 teachers to evaluate their feedback-giving literacy in classroom settings, using Item Response Theory (IRT), Exploratory Factor Analysis (EFA), and Confirmatory Factor Analysis (CFA). The scale was constructed based on a conceptual framework encompassing four components: knowledge, skills, values, and actionability. A pilot test with 1068 teachers led to the selection of 30 items, which were then validated with a sample of 980 teachers. EFA revealed a clear factor structure, explaining 65.42% of the total variance, while CFA confirmed a good model fit (CFI > 0.9, RMSEA < 0.08). The final scale demonstrated high internal consistency (McDonald's Omega coefficient = 0.97) across all subscales. IRT analyses indicated strong measurement precision, particularly in the skills and actionability subscales. Although limited to the Chinese K-12 context and based on self-reported data, the findings offer a valuable tool for teachers to assess and improve their feedback practices. The scale can be used for professional development and further research on feedback-giving literacy. Future studies should explore its applicability in different cultural contexts and investigate the development of teacher feedback literacy over time.

<sup>1</sup> The School of Teacher Education, Jiangsu University, Zhenjiang, China. <sup>2</sup> The Evidence-based Research Center for Educational Assessment (ERCEA), Jiangsu University, Zhenjiang, China. <sup>3</sup> These authors contributed equally: Cui Kang, Jinyan Huang. ✉email: [huangniagara@hotmail.com](mailto:huangniagara@hotmail.com)

## Introduction

Feedback is a vital component of the teaching and learning process, playing a crucial role in enhancing student achievement and engagement. Effective feedback practices have long been recognized as essential in improving student learning outcomes (Brown et al., 2012; Hattie & Timperley, 2007; Shute, 2008). Teachers, as the primary feedback providers, influence students' understanding and growth through their ability to offer constructive feedback (Chan and Luo, 2022; Hattie and Gan, 2011; Wiggins, 2012). However, the effectiveness of feedback varies considerably depending on teachers' competencies and self-awareness in delivering it (Boud & Molloy, 2013; Carless and Boud, 2018).

Recent research has highlighted the complexity of effective feedback provision, emphasizing its dependence on various teacher attributes, such as knowledge, skills, and self-efficacy (Evans, 2013; Lee, 2011). Furthermore, teachers' perceptions of their feedback capabilities are closely linked to their actual feedback practices (Britton, 2023; Carless, 2015; Molloy et al., 2020). Despite these findings, there is a noticeable lack of tools that allow teachers to self-assess and reflect on their feedback-giving skills in a structured and reliable way.

A key gap in the current literature is the absence of a standardized, empirically validated self-assessment scale for teachers to evaluate their feedback capabilities in K-12 education (Nicol, Macfarlane-Dick (2006); Panadero et al., 2016). While several frameworks and models exist that describe effective feedback practices, few empirical studies have focused on developing and validating a tool that enables teachers to assess and improve their feedback skills (Brown and Harris, 2014; Jonsson and Panadero, 2018).

Addressing this gap is crucial for multiple reasons. First, a validated self-assessment scale would help teachers identify areas of strength and areas for growth in their feedback practices, ultimately supporting their professional development (Brown and Harris, 2014; Panadero et al., 2016). Second, grounding the scale in item response theory (IRT) and factor analysis ensures that the tool is not only psychometrically reliable but also accurately reflects the key dimensions of effective feedback (Embretson and Reise, 2000; DeVellis, 2017). These techniques are necessary to ensure the scale's validity, allowing it to capture the complexity of feedback provision and produce reliable results across diverse educational contexts.

This research aims to develop and validate such a self-assessment scale, using IRT and factor analysis to ensure its psychometric robustness and applicability in different educational settings. By filling this gap, the study seeks to provide educators with a valuable tool to enhance their feedback practices, thereby contributing to the broader goal of improving teaching quality and fostering better student learning experiences.

## Review of the literature on teacher feedback literacy

**The concept of teacher feedback literacy.** Teacher feedback literacy refers to educators' ability to effectively provide, interpret, and act upon feedback in educational settings. It encompasses knowledge, competencies, and attitudes necessary for designing feedback mechanisms that promote student engagement with feedback and develop students' feedback literacy (Carless and Winstone, 2020; Tai et al., 2021). Tai et al. (2021), using practice architecture theory, highlight the interconnectedness of teacher and learner feedback practices, underlining the collaborative nature of feedback. Despite varied definitions, the core components of teacher feedback literacy include knowledge of feedback, skills in designing feedback processes, and pedagogical values that support feedback use.

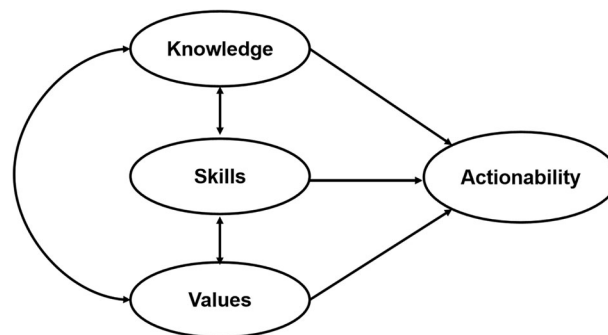
The concept of feedback literacy is widely recognized in educational research, emphasizing the essential role of teachers in fostering environments where students can effectively use feedback. This includes providing guidance and examples to help students implement feedback-driven actions (Carless and Boud, 2018; Winstone and Carless, 2020). While early literature primarily focused on students' feedback literacy in higher education (Carless and Boud, 2018; Han and Xu, 2020; Jonsson, 2013), there has been a growing interest in teacher feedback literacy (Boud and Dawson, 2021; Carless and Winstone, 2020; Heron et al., 2023). Some studies explore feedback literacy in both students and teachers, emphasizing the interaction between these groups in the feedback process (Chen and Liu, 2024; De Kleijn et al., 2019; Guo and Wei, 2019).

**The measurement of teacher feedback literacy.** As the concept of teacher feedback literacy has developed, measuring this construct has gained more attention. Several studies have used qualitative approaches to explore teacher feedback literacy (Boud and Dawson, 2021; Carless and Winstone, 2020; Li and Ke, 2023), while others have taken quantitative approaches to design teacher feedback literacy scales or inventories (Yang et al., 2023; Zhan, 2023). This study aims to contribute to this area by developing an effective scale for assessing teacher feedback literacy.

Previous efforts to develop teacher feedback literacy scales have focused on the reliability and construct validity of proposed models. Zhan (2023) created a feedback literacy scale for pre-service teachers, using confirmatory factor analysis (CFA) to identify six dimensions: planning, implementation, follow-up, appreciation, readiness, and commitment to feedback. Lee et al. (2023) validated a feedback literacy scale for L2 writing teachers, identifying three dimensions: perceived knowledge, values, and perceived skills.

Current instruments for measuring teacher feedback literacy largely rely on classical test theory (CTT). While CTT is widely applicable, it has limitations in areas such as sample size and item constraints (Crocker and Algina, 2006). In contrast, item response theory (IRT) provides a more advanced method for assessing individual abilities and supports adaptive testing, making it a more appropriate tool for large-scale measurement (Embretson and Reise, 2000). This study used IRT to assess the validity and reliability of the scale.

**Framework of teacher feedback-giving literacy.** Teacher feedback-giving literacy refers to educators' ability to effectively design and deliver feedback. As Fig. 1 shows, this framework includes knowledge, skills, values, and actionability — key components essential for providing high-quality feedback that enhances student learning (Carless and Winstone, 2020; Tai et al., 2021; Winstone and Carless, 2020).



**Fig. 1** The framework of teacher feedback-giving literacy.

**Knowledge.** Understanding the principles and practices of effective feedback, including different types and purposes of feedback, and how feedback aligns with learning objectives and students' needs (Hattie and Timperley, 2007; Shute, 2008).

**Skills.** The ability to craft clear, relevant, and constructive feedback, engage students in feedback processes, and provide support for understanding and using feedback (Nicol, Macfarlane-Dick (2006); Wiggins, 2012).

**Values.** Teachers' beliefs and attitudes about feedback, viewing it as a tool for learning rather than just evaluation. This includes recognizing feedback's emotional and motivational impact (Carless and Boud, 2018; Hargreaves, 2013).

**Actionability.** The clarity and supportiveness of feedback, enabling students to take constructive actions toward improvement (Black and Wiliam, 1998; Carless et al., 2011).

## Research aims

This study aims to construct and validate a self-assessment scale for Chinese K-12 teachers to evaluate their feedback-giving literacy in classroom settings. Specifically, the study seeks to: a) develop a generic self-assessment scale for measuring Chinese K-12 teachers' feedback-giving literacy; and b) evaluate the psychometric properties (validity and reliability) of the scale.

## Method

This study was conducted in two phases: the scale construction phase and the scale validation phase. The methodology for each phase is described as follows.

**The scale construction phase.** The scale was crafted within the IRT framework, encompassing the following six major steps:

a) *Crafting item specifications:* Aligned with the teacher feedback-giving literacy conceptual framework (see Fig. 1), item specifications were crafted. Each item was designed to assess Chinese K-12 teachers' proficiency in feedback-giving literacy across the four components: knowledge, skills, values, and actionability. Items were designed to require understanding and recollection of relevant facts and terminology, supported by literature on teacher feedback literacy. Items began with phrases like "I understand," "I am aware of," "I possess the ability," "I value," and "My feedback," to ensure they functioned effectively as self-assessment tools. Additionally, the distribution of items was balanced across the four components.

b) *Assembling the initial pool of scale items:* The first two authors, who are experts in K-12 educational assessment with extensive experience in instrument construction and validation, were responsible for drafting the initial items. Each author individually crafted approximately 50 items, covering all four components. The authors then reviewed the items collectively, selecting 48 valid items, with 12 items allocated to each component, for inclusion in the initial item pool.

c) *Piloting scale items with Chinese K-12 teachers:* The 48-item pool was field-tested with 1,081 Chinese K-12 teachers from various schools across the country. After excluding incomplete responses, 1068 valid responses were analyzed, with an effective response rate of 98.80%. The sample included 269 male (25.2%) and 799 female (74.8%) teachers. Educational qualifications were diverse, including Bachelor's, Master's, and Doctoral degrees. Teachers' occupational status ranged from pre-service to in-service, with representation from urban and rural teaching locations, and across different educational levels.

d) *Conducting EFA to explore the factorial structure:* Exploratory factor analysis (EFA) was performed on the pilot test data to explore the underlying structure of the scale and identify suitable items for the final scale. A maximum likelihood EFA with varimax rotation was used to distinguish the unique factors within the 48 items.

e) *Evaluating item functioning:* The item functioning values were computed using the MULTILOG software. The graded response model (Samejima, 1969) was applied, with one discrimination parameter and four threshold parameters estimated for each item using maximum likelihood estimation. The item information function values were then calculated, representing each item's contribution to the scale's overall information.

f) *Selecting items for the final scale:* Items for the final scale were selected based on two criteria: (1) higher item information function values (e.g., >0.25), and (2) alignment with the typical format of large-scale standardized assessments, which often include around 30 items (Gierl, 2004). Eighteen items with the lowest information function values were excluded, leaving 30 items in the final scale (7, 8, 8, and 7 items for the components of feedback knowledge, feedback skills, feedback values, and feedback actionability, respectively).

**The scale validation phase.** Phase 2 involved validating the scale using IRT, EFA, and CFA frameworks, following these five steps:

a) *Expert evaluation of item appropriateness for construct interpretation:* Following the guidelines of the American Educational Research Association (AERA) et al. (2014), 24 college faculty members with expertise in assessment evaluated the appropriateness of each item for construct interpretation. They used a 1–5 point scale to rate the items in the context of Chinese K-12 teachers' feedback-giving literacy.

b) *Administering the final scale to Chinese K-12 teachers:* The final 30-item scale was administered to 980 Chinese K-12 teachers from various schools across the country. The sample included a diverse representation of gender, educational background, professional status, teaching location, educational level, and teaching subjects.

c) *Conducting EFA and CFA analyses to assess validity:* To adhere to best practices, the dataset was randomly split into two subsamples. EFA was performed on the first subsample ( $N = 488$ ) to assess construct validity, while CFA was conducted on the second subsample ( $N = 492$ ) using AMOS to confirm the factor structure.

d) *Calculating alpha coefficients to measure reliability:* Internal consistency reliability (i.e., McDonald's Omega) coefficients were calculated for the entire scale and for each of the four subscales (i.e., Knowledge, Skills, Values, and Actionability). McDonald's Omega has been shown to provide more accurate reliability estimates than Cronbach's alpha (McNeish, 2018), especially for scales with multidimensional constructs.

e) *Computing item, scale, and subscale information function values to evaluate reliability:* The final scale data were analyzed using MULTILOG to compute item and scale information function values. These values were used to assess the reliability and precision of the scale in measuring feedback-giving literacy among the participants.

## Results

**Expert evaluation of item appropriateness for construct interpretation.** The panel consisted of 24 professionals specializing in teacher education and assessment research. They were tasked with evaluating the appropriateness of each item on the final scale for construct interpretation using a 1–5 point scale (1 =

not appropriate at all; 5 = highly appropriate). The evaluation aimed to assess the suitability of the items for Chinese K-12 teachers to self-assess their feedback-giving literacy. The descriptive statistics are presented in Table 1. As shown in Table 2, all 30 items received a mean score above 4, indicating that the experts deemed these items highly appropriate for the intended construct interpretations.

Table 1 A brief summary the demographic and professional details of the expert panel.		
Category	Detail	N
Gender	Male	12
	Female	12
Academic Rank	Senior Professors	9
	Associate Professors	9
	Intermediate Level	6
Educational Qualification	Doctoral Degrees	23
	Master's Degree	1
Major	Social Sciences	17
	Natural Sciences	3
	Humanities	4
Teaching Experience	1 to 5 years	7
	6 to 10 years	1
	11 to 15 years	2
	16 to 20 years	9
	21 to 25 years	4
	26 to 30 years	1

**EFA results.** EFA was conducted using SPSS with a sample of 488 participants, comprising 106 males and 381 females, with one missing response. The extraction method used was maximum likelihood, and the rotation method employed was varimax. The analysis revealed that all factor loadings exceeded .4, and the communalities for the items were also above .4 (see Table 3), indicating that the factors were adequately represented. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .963, and Bartlett's test of sphericity was significant ( $\chi^2 = 12,652.095$ ,  $df = 435$ ,  $p < 0.001$ ), suggesting that the data were suitable for factor analysis. The cumulative variance explained by the extracted factors was 65.42%, demonstrating that the model accounted for a substantial proportion of the variance in the data.

**CFA results.** CFA was performed using AMOS with a sample size of 492 participants. The fit indices for the model were as follows (Table 4): the chi-square to degrees of freedom ratio was 2.839, which is below the threshold of 3, indicating a good fit. The Comparative Fit Index (CFI), Normed Fit Index (NFI), Incremental Fit Index (IFI), and Tucker-Lewis Index (TLI) were all greater than .9, reflecting a strong model fit. The Standardized Root Mean Square Residual (SRMR) was .0747, and the Root Mean Square Error of Approximation (RMSEA) was 0.061<sup>1</sup>, suggesting a well-fitting model (Browne and Cudeck, 1993; Hu and Bentler, 1999). The factor loadings for the items ranged from 0.48 to 0.91, with all values exceeding the 0.4 threshold (Stevens, 1992), indicating that the items had substantial and reliable loadings on their respective factors. Figure 2 is the CFA model of teacher feedback literacy scale.

Table 2 Expert evaluation of item appropriateness for construct interpretation.		
Items	Appropriateness	
	Mean	SD
K1: I understand the theoretical principles that guide effective feedback.	4.38	0.77
K2: I understand different feedback strategies and their application.	4.54	0.66
K3: I am concerned with feeding back relevant research to enhance my teaching practice.	4.21	1.10
K4: I am familiar with different types of feedback and the purposes for which they are used.	4.38	0.65
K5: I am aware of the impact of timely and appropriate feedback on student learning outcomes.	4.67	0.64
K6: I know the importance of feedback being aligned with student learning goals and needs.	4.63	0.65
K7: I know the importance of developing actionable feedback.	4.58	0.72
S1: I can develop individualized feedback based on student needs.	4.50	0.72
S2: I am able to use positive feedback and constructive criticism wisely.	4.33	0.92
S3: I was able to actively engage students in the feedback process.	4.50	0.59
S4: I was able to convene students to form study groups to conduct a student-to-student assessment.	4.33	0.76
S5: I am able to use feedback to help students learn new knowledge and skills.	4.46	0.72
S6: I am able to design feedback to help students identify gaps and fill in gaps.	4.63	0.58
S7: I am able to ensure that my feedback helps students achieve their learning goals.	4.33	0.70
S8: I am able to ensure that the feedback I provide is viable and allows students to learn through improvement.	4.54	0.59
V1: I don't see feedback as making simple judgments, but as an opportunity for personal growth and development for teachers.	4.71	0.55
V2: I believe that constructive feedback is conducive to creating a positive learning environment.	4.71	0.55
V3: I believe that feedback promotes student autonomy and self-regulated learning.	4.67	0.57
V4: I believe that engaging students in feedback enhances their motivation to learn.	4.71	0.55
V5: I think feedback has an impact on students' emotions.	4.33	0.82
V6: I believe that feedback challenges and grows students.	4.67	0.57
V7: When providing constructive feedback, I think it is important to use appropriate language and tone.	4.54	0.72
V8: I think building a trusting relationship with students makes them more open to feedback.	4.63	0.65
A1: My feedback is easy to understand and implement.	4.46	0.59
A2: My feedback takes into account the student's strengths, but also emphasizes specific areas for improvement.	4.50	0.72
A3: My feedback encourages students to reflect and improve.	4.63	0.50
A4: My feedback includes specific, actionable steps that are easy for students to understand and implement.	4.54	0.59
A5: My feedback helps students understand their current level of learning and how to improve in the future.	4.50	0.66
A6: I can offer actionable feedback at the right time.	4.54	0.66
A7: My feedback motivates students to go deeper into the course content.	4.67	0.48

Note: SD = standard deviation.



**Table 3 EFA results (Factor loadings and communalities).**

Items	F1	F2	F3	F4	CM
K1: I understand the theoretical principles that guide effective feedback.	0.68				0.56
K2: I understand different feedback strategies and their application.	0.74				0.69
K3: I am concerned with feeding back relevant research to enhance my teaching practice.	0.68				0.59
K4: I am familiar with different types of feedback and the purposes for which they are used.	0.66				0.62
K5: I am aware of the impact of timely and appropriate feedback on student learning outcomes.	0.54				0.54
K6: I know the importance of feedback being aligned with student learning goals and needs.	0.48				0.48
K7: I know the importance of developing actionable feedback.	0.43				0.47
S1: I can develop individualized feedback based on student needs.		0.49			0.61
S2: I am able to use positive feedback and constructive criticism wisely.		0.52			0.60
S3: I was able to actively engage students in the feedback process.		.62			0.71
S4: I was able to convene students to form study groups to conduct a student-to-student assessment.		0.60			0.61
S5: I am able to use feedback to help students learn new knowledge and skills.		0.65			0.74
S6: I am able to design feedback to help students identify gaps and fill in gaps.		0.63			0.72
S7: I am able to ensure that my feedback helps students achieve their learning goals.		0.62			0.71
S8: I am able to ensure that the feedback I provide is viable and allows students to learn through improvement.		0.59			0.70
V1: I don't see feedback as making simple judgments, but as an opportunity for personal growth and development for teachers.			0.73		0.65
V2: I believe that constructive feedback is conducive to creating a positive learning environment.			0.80		0.73
V3: I believe that feedback promotes student autonomy and self-regulated learning.			0.84		0.80
V4: I believe that engaging students in feedback enhances their motivation to learn.			0.79		0.70
V5: I think feedback has an impact on students' emotions.			0.61		0.44
V6: I believe that feedback challenges and grows students.			0.69		0.59
V7: When providing constructive feedback, I think it is important to use appropriate language and tone.			0.71		0.60
V8: I think building a trusting relationship with students makes them more open to feedback.			0.75		0.64
A1: My feedback is easy to understand and implement.				0.66	0.70
A2: My feedback takes into account the student's strengths, but also emphasizes specific areas for improvement.				0.68	0.73
A3: My feedback encourages students to reflect and improve.				0.69	0.75
A4: My feedback includes specific, actionable steps that are easy for students to understand and implement.				0.69	0.75
A5: My feedback helps students understand their current level of learning and how to improve in the future.				0.72	0.75
A6: I can offer actionable feedback at the right time.				0.70	0.75
A7: My feedback motivates students to go deeper into the course content.				0.71	0.73

F1 Factor 1 (feedback knowledge), F2 Factor 2 (feedback skills), F3 Factor 3 (feedback values), F4 Factor 4 (feedback actionability), CM communalities.

**Table 4 Model fit indices of teacher feedback literacy scale.**

CMIN	DF	CMIN/DF	RMSEA	CFI	NFI	IFI	TLI	SRMR
1124.112	396	2.839	0.061	0.940	0.911	0.940	0.934	0.0747

**McDonald's Omega coefficients (reliability).** As shown in Table 5, the reliability of the scale and its subscales was assessed using McDonald's Omega coefficients (McNeish, 2018). The subscales demonstrated the following Omega values: feedback knowledge (FK) with 7 items had an Omega of .89, feedback skills (FS) with 8 items and feedback values (FV) with 8 items both had Omegas of 0.95, and feedback actionability (FA) with 7 items had an Omega of .96. The overall 30-item scale demonstrated high reliability with an Omega of .97. These results indicate excellent internal consistency across the subscales and the total scale.

### IRT results

Item Response Theory (IRT) analysis was conducted using MULTILOG with the 30-item scale. The item and scale information functions are summarized in Tables 6 and 7, respectively. These tables present the information provided by individual items, subscales, and the overall scale across the range of the latent trait. The results indicated that the scale provides substantial measurement precision across the trait continuum, as evidenced by the information functions.

Table 6 presents the IRT information function values for the 30 final scale items, ranging from −3 to +3 on the ability scale. For

Feedback Knowledge, items #K1 through #K7 exhibit moderate to high levels of information, with item #K5 showing the highest value of 0.85 at −3. The items for Feedback Skills, including #S1 through #S8, demonstrate consistently high information values, with items #S5 and #S6 reaching peaks of 2.61 and 2.07, respectively. For Feedback Values, items #V1 through #V8 display variable information values, with items #V2 and #V3 achieving the highest values of 1.04 and 1.07, respectively. Lastly, for Feedback Actionability, items #A1 through #A7 maintain high information levels, with item #A5 reaching the maximum value of 2.66 at −3. It should be noted that 28 items at +3 have information function values lower than .25, and item #V8 has information function values lower than 0.25 at +2. Overall, the Feedback Skills and Feedback Actionability subscales are the most informative across the range of ability levels, indicating their strong measurement properties.

Table 7 presents the IRT information function (IF) and standard error (SE) values for the final scale and its subscales across various ability levels. The final scale shows peak information at −2 with an IF value of 43.54 and the highest SE at +3 with 0.48. For the Feedback Knowledge subscale, information is highest at −2 (IF = 11.76), with the greatest SE at +3 (0.91). The Feedback Skills subscale reaches its maximum information value of 26.17 at

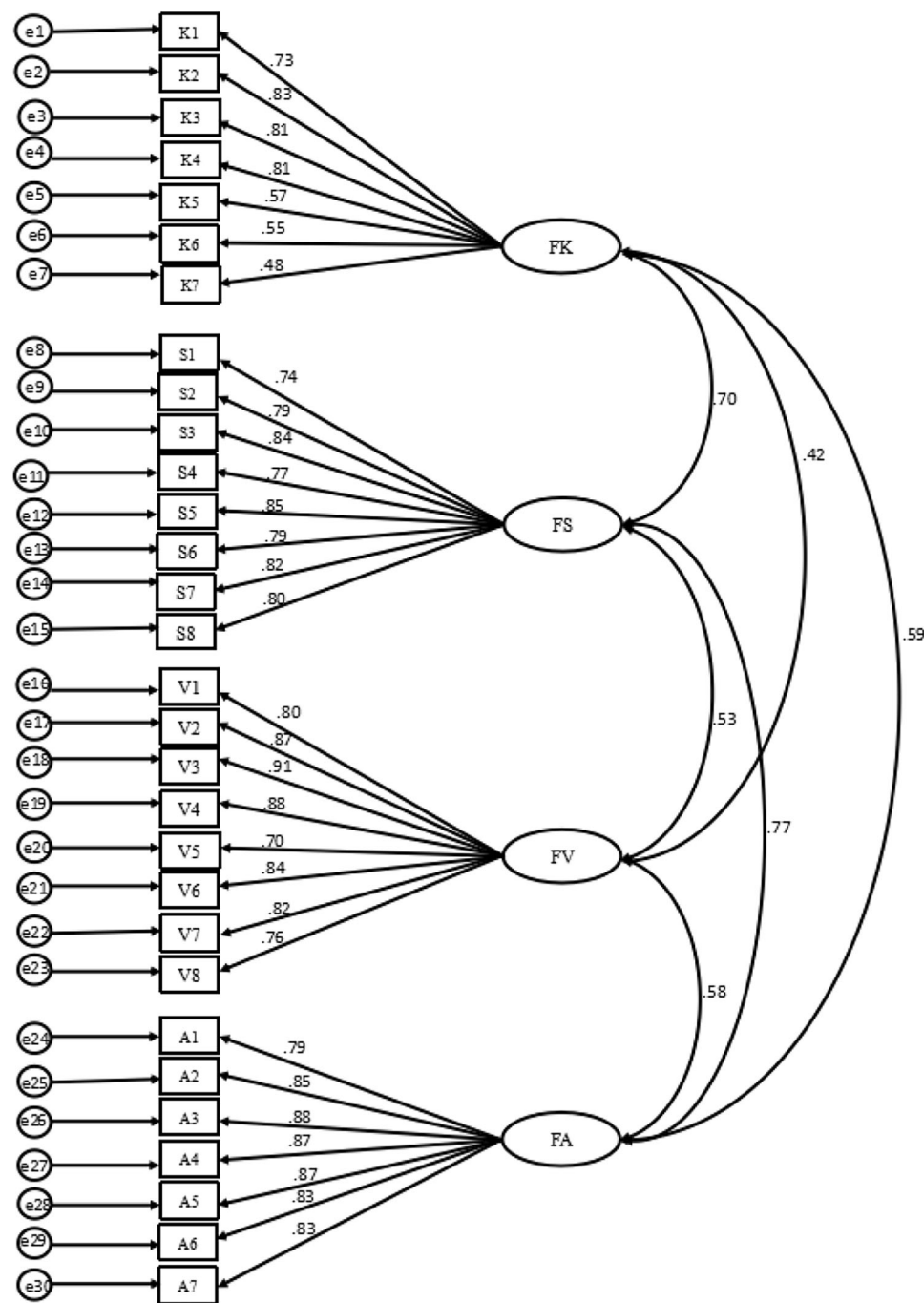


Fig. 2 The CFA model of teacher feedback literacy scale.

Table 5 McDonald's Omega reliability coefficients of the overall scale and each subscale.		
Subscale	Number of items	Omega
FK	7	0.89
FS	8	0.95
FV	8	0.95
FA	7	0.96
Overall scale	30	0.97

FK feedback knowledge, FS feedback skills, FV feedback values, FA feedback actionability.

−2, while the SE is highest at +3 (0.94). The Feedback Values subscale has its highest information at −2 (IF = 30.02) and the largest SE at +3 (0.99). Finally, the Feedback Actionability (FA) subscale exhibits its peak information of 31.38 at −2, with the highest SE at +3 (0.99). This data indicates that the subscales provide the most information at lower ability levels, while the standard errors increase at higher ability levels.

Discussion and conclusions

Discussion of the results. The findings of this study provide significant contributions to the understanding and measurement of teacher feedback-giving literacy, particularly within the context

**Table 6 The information function values of the 30 final scale items.**

Item	-3	-2	-1	0	1	2	3
K1	0.51	0.53	0.49	0.51	0.50	0.39	<b>0.17</b>
K2	0.58	0.66	0.61	0.62	0.60	0.59	0.31
K3	0.67	0.74	0.66	0.70	0.68	0.54	<b>0.19</b>
K4	0.63	0.75	0.68	0.69	0.68	0.69	0.38
K5	0.85	0.78	0.72	0.77	0.78	0.48	<b>0.13</b>
K6	0.77	0.64	0.70	0.64	0.69	0.38	<b>0.10</b>
K7	0.81	0.70	0.74	0.70	0.73	0.37	<b>0.09</b>
S1	0.87	1.60	1.34	1.49	1.11	1.19	<b>0.21</b>
S2	1.51	1.83	1.18	1.60	1.27	1.14	<b>0.14</b>
S3	1.64	2.12	1.41	2.12	1.59	1.21	<b>0.09</b>
S4	1.03	1.44	1.10	1.29	1.11	0.94	<b>0.16</b>
S5	0.94	2.61	1.71	2.13	1.74	1.14	<b>0.07</b>
S6	0.92	2.07	1.52	1.72	1.44	1.07	<b>0.11</b>
S7	1.24	1.84	1.53	1.83	1.42	1.15	<b>0.12</b>
S8	1.42	2.18	1.52	2.03	1.49	1.27	<b>0.10</b>
V1	0.86	0.75	0.79	0.67	0.75	0.37	<b>0.09</b>
V2	1.04	0.88	1.02	0.85	0.98	0.35	<b>0.06</b>
V3	1.07	0.86	0.95	0.77	0.91	0.36	<b>0.07</b>
V4	0.90	0.62	0.90	0.75	0.87	0.35	<b>0.07</b>
V5	0.56	0.52	0.54	0.49	0.51	0.34	<b>0.12</b>
V6	0.93	0.73	0.88	0.72	0.85	0.36	<b>0.07</b>
V7	0.86	0.74	0.79	0.69	0.71	0.27	<b>0.06</b>
V8	0.74	0.66	0.68	0.63	0.62	<b>0.24</b>	<b>0.06</b>
A1	1.27	1.95	1.26	1.63	1.39	1.05	<b>0.11</b>
A2	2.08	2.30	1.47	1.84	1.88	0.90	<b>0.06</b>
A3	2.20	2.42	1.80	2.00	2.18	0.87	<b>0.04</b>
A4	2.26	2.87	1.62	2.35	2.49	0.82	<b>0.03</b>
A5	2.66	2.53	1.44	2.16	2.23	0.83	<b>0.04</b>
A6	2.52	2.06	1.55	2.31	1.92	1.06	<b>0.06</b>
A7	2.39	2.19	1.47	2.04	1.98	0.90	<b>0.05</b>

The value in bold is less than 0.25.

**Table 7 Scale and subscale information and standard error values.**

Scale	Value	-3	-2	-1	0	1	2	3
Final	IF	37.72	43.54	34.05	39.72	37.09	22.59	4.34
scale	SE	0.16	0.15	0.17	0.16	0.16	0.21	0.48
FK	IF	9.74	11.76	10.18	10.76	10.14	3.07	1.20
	SE	0.32	0.29	0.31	0.31	0.31	0.57	0.91
FS	IF	19.39	26.17	17.70	13.38	24.75	3.78	1.13
	SE	0.23	0.20	0.24	0.27	0.20	0.51	0.94
FV	IF	20.23	30.02	33.29	17.20	10.78	1.45	1.03
	SE	0.22	0.18	0.17	0.24	0.31	0.83	0.99
FA	IF	24.89	31.38	20.76	10.80	28.68	2.15	1.03
	SE	0.20	0.18	0.22	0.30	0.19	0.68	0.99

IF information function, SE standard error, FK feedback knowledge, FS feedback skills, FV feedback value, FA feedback actionability.

of Chinese K-12 education. The construction and validation of the self-assessment scale, underpinned by robust methodologies such as IRT, EFA, and CFA, offer new insights into the psychometric properties and practical applicability of such tools in diverse educational settings. The scale's reliability and validity, as evidenced by high McDonald's Omega coefficients and strong factor loadings, extend and complement the existing literature on feedback literacy.

The high reliability and robust psychometric properties of the scale are consistent with prior studies that emphasized the need for reliable and valid measures of teacher feedback literacy (Zhan, 2023; Yang et al., 2023). Similar to Zhan (2023), who developed a

feedback literacy scale for pre-service teachers, this study confirms that teacher feedback literacy can be effectively captured through well-designed scales. The identified components — feedback knowledge, skills, values, and actionability — align with frameworks proposed by Carless and Winstone (2020), highlighting the importance of a comprehensive understanding of feedback mechanisms in education. The balanced distribution of items across these components reflects the multidimensional nature of feedback literacy, as conceptualized in the literature.

Moreover, the strong fit indices observed in the CFA, such as CFI and RMSEA, suggest that the scale's factor structure is both robust and consistent with theoretical expectations. These findings align with previous research by Lee et al. (2023), who used EFA and CFA to validate a feedback literacy scale for L2 writing teachers. The substantial factor loadings observed across the items further demonstrate that the scale reliably measures distinct but related aspects of feedback-giving literacy, echoing the theoretical foundations described by Tai et al. (2021).

This study provides novel insights that contrast with some existing research. While earlier studies have often focused on specific teacher subgroups, such as pre-service or L2 writing teachers (Lee et al., 2023), this study's broad application to a diverse sample of Chinese K-12 teachers enhances the generalizability of the findings. The scale's applicability across various educational contexts, as demonstrated by its strong psychometric properties in different teacher subgroups, suggests that the components of feedback literacy may be universally relevant. This extends the understanding of feedback literacy beyond the specific contexts explored in previous studies.

The findings related to the item information function also highlight the scale's precision in measuring feedback-giving literacy across different levels of ability. The high information values associated with the feedback skills and actionability subscales particularly highlight the critical role these components play in effective feedback practices. This aligns with the emphasis on actionable feedback in the literature (Black and Wiliam, 1998; Carless, 2023; Carless et al., 2011), suggesting that teachers' ability to provide feedback that is clear, understandable, and supportive is crucial for fostering student learning.

In conclusion, this study not only validates a new tool for assessing teacher feedback-giving literacy but also contributes to the broader discourse on feedback literacy by confirming the multidimensional nature of the construct and its relevance across diverse educational settings, particularly within the context of Chinese K-12 education. By examining a wide range of teachers across different school types and grade levels, the findings highlight the universal applicability of the feedback literacy components and the importance of developing reliable and valid assessment tools to enhance teachers' self-awareness and professional growth in feedback practices. Future research could explore the longitudinal impact of using such scales on teachers' feedback practices and student outcomes, further expanding the understanding of feedback literacy in educational contexts.

**Limitations of the study.** While this study contributes valuable insights into the development and validation of a self-assessment scale for teacher feedback-giving literacy, several limitations must be acknowledged. First, the study's sample was limited to Chinese K-12 teachers, which may restrict the generalizability of the findings to other educational contexts or cultural settings. Although the scale demonstrated strong psychometric properties within this sample, the specific educational practices and cultural norms in China may influence how feedback-giving literacy is perceived and practiced, potentially limiting the applicability of the scale in different countries or regions. Given that feedback-

giving literacy may be influenced by contextual factors, future research should explore how feedback literacy is perceived and implemented in other educational settings to assess whether the scale remains effective across diverse cultural and educational contexts.

Second, the study relied on self-reported data for both the initial item pool and the final scale validation. While self-assessment is a valuable tool for measuring teachers' perceptions of their feedback-giving literacy, it may be subject to biases such as social desirability or lack of self-awareness (Brenner and DeLamater, 2016). Teachers may overestimate their skills, knowledge, or practices, especially when they perceive the assessment as a reflection of their professional competence, or in contexts where there may be pressure to appear competent (Hattie and Gan, 2011). While advanced statistical methods such as IRT were used to ensure measurement precision, the potential for bias in self-reports still exists and may affect the validity of the findings.

Third, the cross-sectional design of the study means that it captures feedback-giving literacy at a single point in time, without accounting for potential changes in teachers' abilities or practices over time. Longitudinal studies would be needed to explore how teachers' feedback-giving literacy evolves with experience, professional development, or changes in educational policies. Furthermore, the self-assessment tool may need further validation over time, particularly to assess whether the tool accurately reflects changes in teachers' feedback practices and professional growth.

Finally, the tool does not include external or observational data to cross-check the self-reports. Without independent verification mechanisms, the tool fails to provide an entirely accurate picture of teachers' true feedback-giving literacy. Future studies could incorporate peer evaluations, classroom observations, or other objective measures to validate self-reported data, offering a more comprehensive and balanced view of teachers' feedback practices.

## Conclusions

In light of these limitations, three key conclusions can be drawn from the study. First, the study successfully developed and validated a reliable and robust self-assessment scale for evaluating teacher feedback-giving literacy among Chinese K-12 teachers. Despite the study's cultural and contextual limitations, the scale's strong psychometric properties indicate that it is a valuable tool for assessing key components of feedback literacy, such as knowledge, skills, values, and actionability. This tool can help educators reflect on and improve their feedback practices, potentially enhancing teaching effectiveness and student learning outcomes.

Second, the findings reinforce the notion that feedback-giving literacy is a multidimensional construct, encompassing a range of competencies crucial for effective feedback practices. The high reliability and validity of the subscales, particularly in feedback skills and actionability, highlight the importance of these dimensions in the overall feedback process. This supports existing literature that emphasizes the complexity of feedback literacy and the need for comprehensive assessment tools that capture its various facets.

Finally, while the scale shows promise within the Chinese K-12 context, its applicability across different cultural and educational settings remains uncertain. The study's reliance on self-reported data and a cross-sectional design suggests that future research should consider adapting and testing the scale in diverse contexts and over time. Longitudinal studies could provide deeper insights into how feedback-giving literacy develops throughout teachers' careers and how the scale might be refined to better capture these changes.

**Educational implications.** The findings of this study have several specific implications for teachers and researchers in the Chinese K-12 educational context. The validated self-assessment scale provides a practical tool for Chinese K-12 teachers to self-assess their feedback-giving literacy, fostering greater self-awareness of their feedback practices. By using this scale, teachers can identify strengths and areas for improvement in their feedback-giving skills, knowledge, and attitudes. This reflective process can support professional development, enabling teachers to better understand how their feedback impacts student learning and to take targeted actions to enhance their effectiveness.

The study highlights the critical role of actionable, constructive feedback in the learning process. The focus on skills and actionability within the scale highlights the importance of providing feedback that is clear, relevant, and supportive of student growth. Teachers can use insights from their self-assessment to refine their feedback approaches, potentially leading to more engaged students and improved learning outcomes in Chinese K-12 classrooms.

For researchers in the Chinese K-12 context, this study offers a validated framework for exploring teacher feedback literacy on a larger scale. The scale can be employed in future studies to investigate the relationship between teacher feedback literacy and student performance, or to assess the impact of professional development programs aimed at enhancing feedback practices. Moreover, the findings provide a foundation for educational policymakers to design interventions and support systems that promote effective feedback-giving literacy among teachers, contributing to the broader goal of improving teaching quality in the Chinese K-12 system.

**Directions for future research.** While this study makes significant contributions to the field of teacher feedback literacy, five areas warrant further exploration to enhance the tool's effectiveness and broaden its impact. First, future research should aim to integrate objective assessments, peer evaluations, or classroom observations with the self-assessment tool. This would help mitigate self-report biases and provide a more holistic view of teachers' feedback-giving literacy. Using multiple data sources could allow for the triangulation of findings, enhancing the credibility of the assessment and its practical applicability. Peer evaluations and direct observations could offer invaluable insights that complement self-reports, offering more reliable measures of teaching effectiveness.

Second, future studies should adopt a longitudinal design to track how feedback-giving literacy evolves over time. This would allow researchers to examine how teachers' feedback practices develop as they gain more experience, engage in professional development programs, or adapt to changes in educational policies. Longitudinal studies could provide a clearer understanding of the tool's role in supporting ongoing professional growth and how teachers' self-reported literacy relates to actual improvements in their teaching practices.

Third, it would be valuable to examine the relationship between feedback-giving literacy and student outcomes. Future research could investigate whether improvements in teachers' self-reported feedback literacy lead to tangible benefits in student performance, engagement, or learning outcomes. This could include correlational studies between teachers' feedback literacy scores and various student success metrics, providing evidence for the impact of feedback-giving practices on educational results.

Fourth, given the limitations of the study's context, further research should explore the applicability of the feedback-giving literacy scale in different educational and cultural settings. By adapting the tool to various regions, researchers can assess



whether the components of feedback literacy remain relevant across diverse educational systems. This would allow for a better understanding of how cultural and educational contexts influence feedback-giving practices and how the tool can be modified for broader use.

Finally, future research could examine how the self-assessment tool can be used in professional development programs to support teachers in improving their feedback-giving practices. Studies could explore the effectiveness of professional development initiatives that incorporate the scale as a reflective tool. Research could assess how teachers' feedback practices change over time as a result of targeted professional development, providing valuable insights into the tool's potential to enhance teaching quality.

By addressing these areas, future research can further refine the scale and expand its potential for improving feedback practices in diverse educational settings. Additionally, it could offer deeper insights into the complex relationship between feedback-giving literacy, teaching effectiveness, and student outcomes.

### Data availability

The participants did not provide written informed consent to share their data publicly. However, they did provide written informed consent to share their data with individuals on reasonable request. Therefore, the datasets generated during the current study are available from the corresponding author ONLY on reasonable request.

Received: 9 September 2024; Accepted: 23 April 2025;

Published online: 06 May 2025

### Note

1 A value below .05 indicates a good fit (Hu and Bentler, 1999), but values up to .08 are generally considered indicative of a reasonable fit, especially with smaller sample size or in complex models (Browne and Cudeck, 1993).

### References

- American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) (2014). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association
- Black P, Wiliam D (1998) Assessment and classroom learning. *Assess Educ: Princ, Policy Pract* 5(1):7–74
- Boud D, Dawson P (2021) What feedback literate teachers do: An empirically-derived competency framework. *Assess Evaluation High Educ* 48(2):158–171
- Boud D, Molloy E (2013) Rethinking models of feedback for learning: The challenge of design. *Assess Evaluation High Educ* 38(6):698–712
- Brenner, PS, & DeLamater, J (2016). Lies, damned lies, and survey self-reports? Identity as a cause of measurement bias. *Social Psychology Quarterly*, 79(4). <https://doi.org/10.1177/0190272516628298>
- Britton ER (2023) Developing teacher feedback literacy through self-study: Exploring written commentary in a critical language writing curriculum. *Assess Writ* 56:100709
- Browne, MW, & Cudeck, R (1993). Alternative ways of assessing model fit. In KA Bollen & JS Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage
- Brown GTL, Harris LR (2014) The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learn Res* 3(1):22–30
- Brown GTL, Harris LR, Harnett J (2012) Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teach Teach Educ* 28(7):968–978
- Carless, D (2015). *Excellence in university assessment: Learning from award-winning practice*. Routledge
- Carless D (2023) Teacher feedback literacy, feedback regimes and iterative change: Towards enhanced value in feedback processes. *High Educ Res Dev* 42(8):1890–1904
- Carless D, Boud D (2018) The development of student feedback literacy: Enabling uptake of feedback. *Assess Evaluation High Educ* 43(8):1315–1325
- Carless D, Salter D, Yang M, Lam J (2011) Developing sustainable feedback practices. *Stud High Educ* 36(4):395–407
- Carless D, Winstone N (2020) Teacher feedback literacy and its interplay with student feedback literacy. *Teach High Educ* 28(1):150–163
- Chan CKY, Luo J (2022) Exploring teacher perceptions of different types of 'feedback practices' in higher education: Implications for teacher feedback literacy. *Assess Evaluation High Educ* 47(1):61–76
- Chen C, Liu AJ (2024) Understanding partnerships in teacher and student feedback literacy: Shared responsibility. *Innov Educ Teach Int* 61(1):31–44. <https://doi.org/10.1080/14703297.2022.2153722>
- Crocker, L, & Algina, J (2006). *Introduction to classical and modern test theory*. Wadsworth Publishing
- De Kleijn R, Mainhard T, Meijer PC, Pilot A, Brekelmans M (2019) Master's thesis supervision: Relations between supervisor feedback, student self-efficacy and performance. *Assess Evaluation High Educ* 44(7):981–995
- DeVellis, RF (2017). *Scale development: Theory and applications (4th Ed.)*. Sage Publications
- Embretson, SE, & Reise, SP (2000). *Item response theory for psychologists*. Psychology Press
- Evans C (2013) Making sense of assessment feedback in higher education. *Rev Educ Res* 83(1):70–120
- Gierl, M (2004). *Item response theory* (class notes). University of Alberta, Edmonton, Canada
- Guo W, Wei J (2019) Teacher feedback and students' self-regulated learning in mathematics: A study of Chinese secondary students. *Asia-Pac Educ Researcher* 28(3):265–275
- Han Y, Xu Y (2020) The development of student feedback literacy: The influences of teacher feedback on peer feedback. *Assess Evaluation High Educ* 45(5):680–696
- Hargreaves E (2013) Inquiring into children's experiences of teacher feedback: Reconceptualising assessment for learning. *Oxf Rev Educ* 39(1):21–35
- Hattie, J, & Gan, M (2011). Instruction based on feedback. In RE Mayer & PA Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 263–285). Routledge
- Hattie J, Timperley H (2007) The power of feedback. *Rev Educ Res* 77(1):81–112
- Heron M, Medland E, Winstone N, Pitt E (2023) Developing the relational in teacher feedback literacy: Exploring feedback talk. *Assess Evaluation High Educ* 48(2):172–185
- Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling* 6(1):1–55
- Jonsson A (2013) Facilitating productive use of feedback in higher education. *Act Learn High Educ* 14(1):63–76
- Jonsson A, Panadero E (2018) The use and nature of rubrics in peer assessment. *Assess Eval High Educ* 43(5):1–17
- Lee I (2011) Feedback revolution: What gets in the way? *ELT J* 65(1):1–12
- Lee I, Karaca M, Inan S (2023) The development and validation of a scale on L2 writing teacher feedback literacy. *Assess Writ* 57:100743
- Li, X, & Ke, P (2023). Becoming feedback literate: A case study of a Chinese university EFL teacher. *Asia-Pacific Educ Res* <https://doi.org/10.1007/s40299-023-00743-z>
- McNeish D (2018) Thanks coefficient alpha, we'll take it from here. *Psychological Methods* 23(3):412–433
- Molloy E, Boud D, Henderson M (2020) Developing a learning-centred framework for feedback literacy. *Assess Eval High Educ* 45(4):527–540
- Nicol DJ, Macfarlane-Dick D (2006) Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Stud High Educ* 31(2):199–218
- Panadero, E, Jonsson, A, & Strijbos, JW (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In L Verschaffel, E De Corte, J Elen, & JGG van Merriënboer (Eds.), *Instructional psychology: Past, present, and future trends* (pp. 143–167). Routledge
- Samejima F (1969) Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monogr Suppl* 34(4, Pt. 2):100
- Shute VJ (2008) Focus on formative feedback. *Rev Educ Res* 78(1):153–189
- Stevens, J (1992). *Applied multivariate statistics for the social sciences (2nd ed.)*. Lawrence Erlbaum Associates
- Tai J, Bearman M, Gravett K, Molloy E (2021) Exploring the notion of teacher feedback literacies through the theory of practice architectures. *Assess Eval High Educ* 48(2):201–213
- Wiggins G (2012) Seven keys to effective feedback. *Educ Leadersh* 70(1):10–16
- Winstone, N, & Carless, D (2020). *Designing effective feedback processes in higher education: A learning-focused approach*. Routledge
- Yang Y, Luo Z, Dong Y, Kurup PM, Wang Y (2023) Towards a new paradigm: The development and validation of a scale to explore technology-enhanced feedback literacy among primary and secondary school teachers. *Educ Technol Res Dev* 71(2):391–413

Zhan, Y (2023). Feedback literacy of teacher candidates: Roles of assessment course learning experience and motivations for becoming a teacher. *Asia-Pacific Education Researcher*. <https://doi.org/10.1007/s40299-023-00779-1>

## Acknowledgements

This work was supported by (1) the Research on Normal School Students' Assessment Literacy Promotion in Comprehensive University, Jiangsu University Key Project of Higher Educational Reform (2019JGZD016) and (2) Research on the Current Situation and Strategies of "Revitalizing Liberal Arts": Taking Liberal Arts Normal Major in Jiangsu University, Jiangsu University Higher Education Planning and Development Research Project (G202306) granted to Cui Kang.

## Author contributions

Jinyan Huang and Cui Kang made equal contribution to this article. Cui Kang: conceptualization, literature, methodology, data acquisition, data analysis, revision, and funding. Jinyan Huang: conceptualization, literature, methodology, data analysis, writing, and editing for submission. Ying Liu: conceptualization, literature, methodology, data analysis, and revision. Hua Yin: conceptualization, literature, methodology, and data acquisition.

## Competing interests

The authors declare no competing interests.

## Ethical approval

This study adhered to the ethical guidelines outlined in the Code of Ethics for Research in Education as established by the American Educational Research Association. Ethical approval was obtained from the Research Ethical Review Board of the Evidence-based Research Center for Educational Assessment (ERCEA) at Jiangsu University on December 27, 2023 (Approval Number: ERCEA2312).

## Informed consent

The participants had provided their written informed consent before they participated in this study. They have been fully informed the purpose of the study, the voluntary nature of participation and anonymity, how data would be used, and the potential benefits and risks of participating in this study.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1057/s41599-025-04927-4>.

**Correspondence** and requests for materials should be addressed to Jinyan Huang.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025