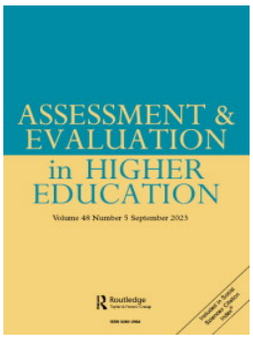

Routledge
Taylor & Francis Group

Assessment & Evaluation in Higher Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/caeh20

Measuring what learners do in feedback: the feedback literacy behaviour scale

Phillip Dawson, Zi Yan, Anastasiya Lipnevich, Joanna Tai, David Boud & Paige Mahoney

To cite this article: Phillip Dawson, Zi Yan, Anastasiya Lipnevich, Joanna Tai, David Boud & Paige Mahoney (2023): Measuring what learners do in feedback: the feedback literacy behaviour scale, Assessment & Evaluation in Higher Education, DOI: 10.1080/02602938.2023.2240983

To link to this article: https://doi.org/10.1080/02602938.2023.2240983

Published online: 07 Aug 2023.

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
https://doi.org/10.1080/02602938.2023.2240983

Measuring what learners do in feedback: the feedback literacy behaviour scale

Phillip Dawson[a], Zi Yan[b], Anastasiya Lipnevich[c], Joanna Tai[a], David Boud[a,d,e] and Paige Mahoney[a]

[a]Centre for Research in Assessment and Digital Learning (CRADLE), Deakin University, Melbourne, Victoria, Australia; [b]Department of Curriculum and Instruction, The Education University of Hong Kong, Hong Kong; [c]City University of New York, Queens College and the Graduate Center New York, NY, USA; [d]Faculty of Arts and Social Sciences, University of Technology Sydney, NSW, Australia; [e]Institute for Work-Based Learning, Middlesex University, London, UK

ABSTRACT

Feedback can be powerful, but its effects are dependent on what students do. There has been intensive research in recent years under the banner of 'feedback literacy' to understand how to help students make the most of feedback. Although there are instruments to measure feedback literacy, they largely measure perceptions and orientations rather than what learners actually do. This paper documents the development and validation of the Feedback Literacy Behaviour Scale (FLBS), which is a self-report instrument intended to measure students' feedback behaviours. A framework for feedback literacy was constructed with five factors: Seek Feedback information (SF), Make Sense of information (MS), Use Feedback information (UF), Provide Feedback information (PF), and Manage Affect (MA). An initial set of 45 questions were reviewed in an iterative process by feedback experts, resulting in 39 questions that were trialled with 350 student participants from four countries. Our final survey of 24 questions was generally supported by confirmatory factor and Rasch analyses, and has acceptable test-retest reliability. The FLBS provides a more robust way for educators and researchers to capture behavioural indicators of feedback literacy and the impact of interventions to improve it.

KEYWORDS

feedback literacy; feedback; assessment; survey; scale development

When feedback works well it has one of the most powerful effects in education (Hattie and Timperley 2007; Wisniewski, Zierer, and Hattie 2019). However, feedback often falls short of its potential, and as a result is one of the most dissatisfying and problematic aspects of education (Boud and Molloy 2013). A common frustration for educators is that the comments they make on student work often go unread or don't lead to change (Winstone, Nash, Rowntree, et al. 2017). In response to the seeming intractability of the problems of feedback, researchers have sought to shift the way we think about feedback, supplementing a focus on the messages teachers provide to students with consideration of how students actively engage in feedback (Carless and Boud 2018). In this view, students are positioned as the primary actors in their own personal feedback processes, making use of inputs from a variety of sources.

One challenge that comes with student-centred feedback models is that they demand much more from students than traditional approaches to feedback. 'New paradigm' (Winstone and Carless 2019) models require students to seek feedback information, actively process it, make

CONTACT Phillip Dawson p.dawson@deakin.edu.au

quality judgements about which inputs will be most useful, and take action. This asks students to be far more active than teacher-centred models of feedback, in which students are often positioned as passive recipients of information.

Feedback literacy has been proposed to conceptualise the capabilities that students need to engage in feedback. The notion has gained significant traction in the higher education literature thanks to conceptual work by Carless and Boud (2018) and Molloy, Boud, and Henderson (2020). Although there are differences between various authors' models, all at their core focus on understanding what students need to do to benefit from feedback processes. Feedback literacy is thought by its proponents to be a worthwhile outcome in and of itself, as well as something that may support students in engaging in feedback and, ultimately, in improving their learning.

Research on feedback literacy is currently blossoming with empirical (Noble et al. 2019; Winstone, Mathlin, and Nash 2019; (e.g. Hoo, Deneen, and Boud 2022) and conceptual (e.g. Gravett 2022; Chong 2021) work. To enable quantitative research into interventions to improve feedback literacy, there is a need for instruments to measure it. Currently six such instruments have undergone a formal validation process: the Student Assessment-Based Feedback Literacy (SAFL) scale (Liao 2021), the Scale of Student Feedback Literacy (SSFL) (Zhan 2022), the Feedback Literacy Scale (Song-FLS) (Song 2022), the L2 Student Writing Feedback Literacy Scale (L2-SWFLS) (Yu, Di Zhang, and Liu 2022), the Feedback Literacy Scale (Yildiz-FLS) (Yildiz, Bozpolat, and Hazar 2022), and the Peer Feedback Literacy Scale (PFLS) (Dong, Gao, and Schunn 2023). Characteristics of these scales are summarised in Table 1. We have not included Winstone, Mathlin, and Nash (2019) study on their *Developing Engagement With Feedback Toolkit*, as their instrument is explicitly disclaimed as "simplistic and exploratory" (p. 3), and they state that their aim "was not to develop and validate such a measure of feedback literacy" (p. 8).

The focus across these six scales is mainly students' beliefs about feedback and their own evaluation of their capabilities to engage in feedback processes. Each scale offers little or no consideration of what students do in feedback processes. However, dominant feedback literacy frameworks emphasise that feedback is an active process, and feedback literacy involves "taking action" (Carless and Boud 2018) and eliciting, processing, enacting and providing feedback information (Molloy, Boud, and Henderson 2020). This suggests that to measure feedback literacy,

Table 1. Features of published instruments to measure feedback literacy.

	Focus of items	Context	Sources and modes of feedback	Includes items about learners providing feedback information
SAFL (Liao 2021)	Beliefs about capabilities; values; understandings; knowledge	Higher education	Teachers, mode agnostic	Yes
SSFL (Zhan 2022)	Beliefs about capabilities; values; understandings; knowledge; willingness	Higher education	Teachers and unspecified others, mode agnostic	No
Song-FLS (Song 2022)	Trust in capabilities of feedback provider; beliefs about importance of feedback and roles	Higher education	Teachers and peers, written feedback	No
L2-SWFLS (Yu, Di Zhang, and Liu 2022)	Beliefs about capabilities; values; understandings; knowledge	L2 student writing	Teachers and unspecified others, mode agnostic	Yes
Yildiz-FLS (Yildiz, Bozpolat, and Hazar 2022)	Beliefs about capabilities; values; affective experience; attitude; adoption of feedback; actions	Higher education	Teachers, mode agnostic	No
PFLS (Dong, Gao, and Schunn 2023)	Beliefs about capabilities; knowledge; attitude; intentions	L2 student writing peer feedback	Peers, mode agnostic	Yes

instruments should focus on what students do – not just what they think they should or can do. A similar approach is taken in other related concepts, such as measuring self-regulated learning with the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich et al. 1993), which mostly consists of items about what students do. Just as with self-regulated learning, the utility of feedback literacy depends on its relation to what students do. Feedback is an active process, and if instruments are to measure feedback literacy as enacted – rather than just students' beliefs and attitudes – then they need to ask students about what they do in feedback processes. The six published instruments do not do this in any great depth.

Although current understandings of feedback literacy in higher education have largely been prompted by Carless and Boud's (2018) paper, there is a long history of research on similar concepts in education, psychology and business, and it is worth exploring instruments to measure those concepts. The extensive work on employees' attitudes towards feedback (Anseel et al. 2015) has some overlap with feedback literacy (Joughin et al. 2021), and has included the development and validation of scales to measure various concepts. One example is the Feedback Orientation Scale (Linderbaum and Levy 2010), which measures receptivity to feedback. While this has been validated and is used within business research contexts, its workplace contextualisation means that it misses some aspects of feedback relevant to higher education and contains language and concepts that are not directly applicable (e.g. "Feedback from supervisors can help me advance in a company"). In addition, it is founded on a fundamentally different set of understandings about the problems of feedback. Although the scale is intended to study the agentic practices of employees around feedback information provided to them, the scale positions feedback as something that is given or done to someone, rather than as an active process driven by the learner. This sort of positioning of feedback in surveys for use in education has been criticised by Winstone, Ajjawi, et al. (2022) and it runs counter to an overall shift in the higher education literature towards viewing feedback as a process rather than an input (Winstone, Boud, et al. 2022). Other scales from this domain (e.g. VandeWalle et al. 2000; Krasman 2010) similarly lack either a full coverage of feedback literacy, are a poor match for the context of higher education, and/or do not position feedback in ways that are compatible with feedback literacy.

There are also a variety of instruments in educational psychology that measure related concepts. For example, Lipnevich and Lopera-Oquendo (2022) developed the Receptivity to Feedback Scale, which comprises experiential and instrumental attitudes towards feedback, and cognitive and behavioural engagement with feedback. Although related, the ideas of receptivity to feedback and feedback literacy have very different conceptual bases. Similarly, the Student Conceptions of Feedback instrument adapted from school education to higher education by Brown, Peterson, and Yao (2016) focuses on students' conceptions of feedback. As discussed above, our understanding of feedback literacy differs from both receptivity to feedback and conceptions of feedback in that we are more interested in behaviours than attitudes.

In the absence of a robust scale specifically focused on feedback literacy behaviours, we set out to develop one.

Methodology

Conceptual framework

We based our conceptualisation of feedback literacy on existing frameworks from the literature that were consistent with an understanding of feedback literacy as behaviour. At the time of developing our instrument, there were two such frameworks. The first of these is Carless and Boud's (2018) conceptual paper, which was the most-cited paper on feedback literacy at the time of writing (>1,000 citations on Google Scholar). This framework is particularly useful because its four components (*appreciating feedback processes; making judgements; managing affect; taking action*) suggest behaviours. The second existing framework we used was Molloy, Boud, and Henderson's (2020)

empirically-based framework, which goes into greater depth about the specific behaviours students are expected to do to be considered feedback literate. The most significant addition in terms of behaviours present in Molloy, Boud and Henderson's framework that is not made explicit in Carless and Boud's framework is the incorporation of learners providing feedback information as part of feedback literacy – the category *acknowledges feedback as a reciprocal process*.

We created a new framework, rather than use an existing framework, as neither was framed entirely within behavioural terms, and neither on its own was as comprehensive as we required. Although both are framed in active terms, and described largely in terms of what students do, they occasionally use verbs like "acknowledges" and "appreciates" which require reframing towards a behavioural focus. To develop our framework from these two existing frameworks we took an iterative process. Firstly, we discussed these and other existing frameworks amongst the project team to identify ways to build a comprehensive behavioural framework for feedback literacy. Then we developed potential frameworks, which were workshopped with the team over three successive iterations. We checked the conceptual frameworks with a critical friend, a senior scholar who has published in the field of feedback literacy, and we also checked our framework with an author of each of the frameworks we were building on. The connection between our framework and the two frameworks we built on is detailed in Figure 1.

Our final conceptual framework consists of five components, with the following definitions:

- *Seek feedback information (SF)*: eliciting feedback information from a variety of sources, including one's own notions of quality and examples of good work;
- *Make sense of information (MS)*: processing, evaluating, and interpreting feedback information;
- *Use feedback information (UF)*: putting feedback information into action to improve the quality of current and/or future work;
- *Provide feedback information (PF)*: considering the work of others and making comments about its quality; and
- *Manage affect (MA)*: persisting in feedback processes despite the emotional challenges they pose.

These five components represent our reworking of the two frameworks into a set of broad behaviours that denote the enactment of feedback literacy.

Instrument development

The instrument development was informed by Hinkin's (1998) guideline. After our conceptual framework was established, an initial set of 45 items was developed by the project team. Consistent with our behavioural focus, we sought to represent each component in a series of items representing behaviours that feedback literate students would undertake in feedback processes. We considered adapting items from other inventories but could not find a substantial set that focused on behaviour. We went through three iterations of the item set at this stage, and discussed the items in a mix of online meetings and asynchronous exchanges. Our initial set of 45 questions was split over the five concepts. The preface to the questions was "please think about what you usually do in your studies, and rate how often you do these things". This set was workshopped by the project team through three iterations before being sent for review to 21 scholars regarded by the authors as international feedback experts (most of whom are named in the acknowledgements of this paper). Fifteen of them rated each question on how essential it was for measuring feedback literacy: essential, useful but not essential, or unnecessary. Experts also commented on any wording issues with questions. Questions that at least 11 experts rated as essential were kept, and others were debated among the project team in terms of how essential they were.

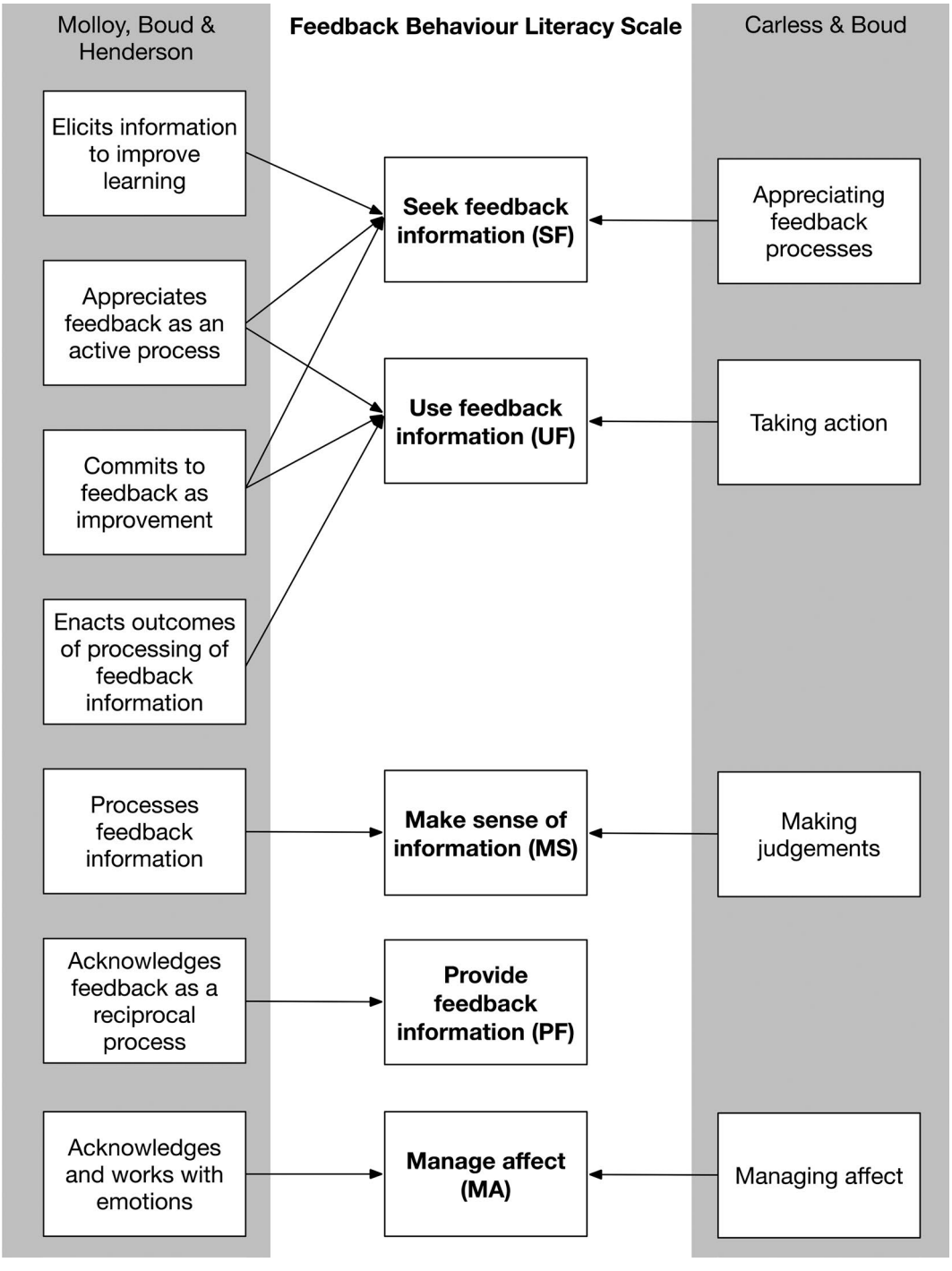

Figure 1. Conceptual framework mapped against components of Molloy, Boud and Henderson's framework and Carless and Boud's framework.

All items in the *provide feedback information* category were rated as not essential by most experts. Analysis of their comments suggested this may have resulted from diverging understandings of feedback literacy between our conceptual framework and the views of the experts.

Our conceptual framework regards learners having the capability to provide feedback information as an essential component of feedback literacy, and this is consistent with the literature; for example, Molloy, Boud, and Henderson (2020) include "Composes useful information for others about the nature of their work" in their list of what feedback literate students do. Three existing scales, the SAFL (Liao 2021), the L2-SWFLS (Yu, Di Zhang, and Liu 2022) and PFLS (Dong, Gao, and Schunn 2023), also include the provision of feedback information as part of feedback literacy. Removing the provision of feedback information by learners from our instrument, or separating it into another instrument, would (a) make our instrument inconsistent with the literature, and (b) represent a missed opportunity to explore relationships between the provision of feedback information and other dimensions of feedback literacy. Because of this, we opted to amend and keep some questions about student provision of feedback information and expose the full set of questions to a second round of review.

This revised set of 39 questions was reviewed by eight experts who again rated every question and provided comments. Five items were dropped based on low scores, and modifications were made to seven other items at this stage, resulting in the instrument of 34 items that was used in this study. The project team, which consists of six feedback researchers from education and psychology, debated items in terms of where they fit in the conceptual framework. A six-point Likert-type frequency response scale (1 = never, 2 = almost never, 3 = rarely, 4 = sometimes, 5 = almost always, 6 = always) was used for all items. A neutral point was excluded to prevent problematic response tendencies (i.e. satisficing, ambivalent responding, etc.) (Johns 2005; Moors 2008). All survey items were mandatory.

The framework for validation

Messick's (1995) framework of validity was adopted to guide the validation. Five out of the six aspects of validity, i.e. content, substantive, structural, generalizability and external aspects, were evaluated in this study. *Content* validity evidence was captured by expert review, as discussed earlier. *Substantive* validity evidence was ensured by the correspondence between items and the underlying conceptual framework (see Figure 1), and the item-level statistics from Rasch analysis, such as the item fit statistics. Confirmatory factor analysis (CFA) results offered evidence of *structural* validity, whereas evidence of *generalizability* could be partially inferred from the results of differential item functioning (DIF) across gender in a Rasch analysis.

To partially understand *external* validity, correlations between subscales of the Big Five Personality Inventory-Short Form (BFI-S) and the FLBS were evaluated. A number of studies have shown that broad personality dimensions of the Big Five personality inventory (extraversion, agreeableness, conscientiousness, emotional stability and openness) often subsume newly introduced concepts (MacCann et al. 2012; Walton et al. 2023). Hence, our goal was to examine links among the subscales of the Feedback Literacy Scale and the Big Five personality factors to ensure construct discrimination and the presence of theoretically meaningful correlations. Prior studies of feedback receptivity showed its links to conscientiousness and openness, with coefficients ranging from $0.164 < r < 0.362$ (see Lipnevich et al. 2021). Personality manifests itself through behaviors, thus, exploring links of the Big Five with factors of FLBS was critical for scale validation (Briley, Domiteaux, and Tucker-Drob 2014). *Consequential* validity could not be reported as we did not have relevant criterion data.

Participants

The instrument was administered to 350 university students (M_{age} = 22.75, SD = 5.54, age range = 18-57 years) recruited through Prolific Academic, which is a paid online participant pool for research surveys. There were 162 (46.3%) males, 181 females (51.7%), and seven participants who identified as transgender or nonbinary. Participants were residing in Australia ($N = 50$, 14.3%),

Ireland ($N = 10$, 2.9%), New Zealand ($N = 15$, 4.3%) and the UK ($N = 275$, 78.6%). Most students ($N = 347$, 99.1%) provided their academic discipline, which we coded as STEM (science, technology, engineering, mathematics: $N = 121$, 34.8%), arts ($N = 83$, 23.9%), health ($N = 78$, 22.4%), business ($N = 42$, 12.1%) or social sciences ($N = 23$, 6.6%).

Data collection

Participants completed the survey online using Qualtrics. All participants (T1) were invited to complete a follow-up survey (T2) with the same questionnaire four weeks later and 322 participants (92%) did. Participants were paid £3.15 for each survey, which was calculated based on the estimated length of time to complete the survey multiplied by the minimum wage in Australia, where the study was based. The data of T1 and T2 were matched through unique identifiers on Prolific Academic. This study was approved by the relevant ethics committee.

In addition to completing the FLBS, participants also completed the *Big Five Personality Inventory-Short Form* (BFI-S), which has 20 items and measures five dimensions of personality (Goldberg 1993). Responses to each item ranged from 1 = strongly disagree to 5 = strongly agree. Previous studies have supported the reliability and validity of the BFI-S with university students (Donnellan et al. 2006; Trapmann et al. 2007).

Data analysis

The psychometric properties of the FLBS were examined *via* both confirmatory factor analysis (CFA) and Rasch analysis. We did not conduct exploratory factor analysis because our scale development was guided by a strong theory. With CFA we explicitly tested a hypothesised model, as discussed earlier, on an empirical data set (Gorsuch 1983). Rasch analysis assesses the quality of a scale by examining the extent to which items in the scale are reflective of a single underlying latent construct. The 'data fit the model' approach adopted in Rasch analysis requires that the collected data meet a priori requirements essential for fundamental measurement purposes (Andrich 2004; Bond, Yan, and Heene 2020). While CFA tests the overall fit between the empirical data and the hypothesised model, Rasch analysis (Rasch 1960) examines the psychometric properties at the item level and, therefore, provides additional evidence regarding the scale quality. Several empirical studies (e.g. Deneen et al. 2013; Testa et al. 2019; Yan et al. 2020) have demonstrated that using both approaches can provide comprehensive and robust evaluation of the quality of an instrument. Rasch analysis provides calibrated scores that account for differences in item difficulty. In this paper, when we report comparisons between an individual's pre-test and post-test performance, and comparisons between the FLBS and the BFI-S, we use these Rasch-calibrated scores. As Rasch analysis is relatively technical and lengthy to describe, and it did not result in the exclusion of any items or changes to our factor structure, we have placed it in supplementary online material.

CFA was conducted using AMOS 24.0 (Arbuckle 2015). Multiple fit indices were checked including the goodness of fit index (GFI), the comparative fit index (CFI), the Tucker-Lewis index (TLI), the standardised root mean square residual (SRMR), and the root mean square error of approximation (RMSEA). GFI, CFI, and TLI with values > 0.90, RMSEA and SRMR with values < 0.08 (Hu and Bentler 1999; McDonald and Ho 2002) indicate an acceptable model-data fit.

The Cronbach's alpha (α; Cronbach 1951) and omega coefficient (ω; McDonald 1999) were computed to evaluate the internal consistency of the scale.

Results

We ran a theory-driven CFA with empirically driven modifications (Mueller and Hancock 2001) with the original five factors: *Seek Feedback information (SF)*, *Make Sense of information (MS)*, *Use*

Feedback information (UF), *Provide Feedback information (PF)*, and *Manage Affect (MA)*. The values for Kaiser–Meyer–Olkin measure of sampling adequacy was 0.907 and Bartlett's test of sphericity was $\chi^2(561) = 4059.532$, $p < 0.001$, indicating that the sample and inter-item correlations were appropriate for factor analysis. The initial results showed that some of the model-data fit statistics were not satisfactory: $\chi^2/df = 2.401$; GFI = 0.814; CFI = 0.801; TLI = 0.784; RMSEA = 0.063; SRMR=.068. Examinations of the factor loadings and modification indices revealed that two items had very low factor loading on the target factor (PF1: *When I comment on someone else's work, I consider the emotional impact of my comments*; MS6: *When making sense of comments I try not to focus on the time and effort I put into the work*), while some items cross loaded on different factors (e.g. MA7: *I am comfortable commenting on other people's work*). According to Saris, Satorra, and Van der Veld (2009) suggestion, we examined the size of modification indices and power of the test for all misfitting items. The results showed misspecification as all modification indices were significant and the power of the test was low. We then removed all questionable items.

A final model was arrived at with five items for each of subscales SF, UF, PF and MA, and four items for the subscale MS. The model-data fit was acceptable: $\chi^2/df = 1.896$; GFI = 0.901; CFI = 0.911; TLI = 0.898; RMSEA = 0.051; SRMR=.053. Figure 2 displays the standardised factor loadings and correlations among the latent traits. The correlation between SF and UF was high (0.94), statistically indicating the need to merge these two factors. However, we chose to keep these factors separate for both conceptual and pragmatic reasons. Firstly, feedback information (which may be obtained through actively seeking feedback or just being provided the information) is received before it is made sense of (MS) and used (UF). Combining SF and UF would mean merging the first and third steps, but not the second, creating conceptually awkward factors. Secondly, we were concerned not to misrepresent student agency in feedback, and to distinguish between situations in which students might not elicit information but are expected to act on teacher comments. Thirdly, we wanted to retain the ability of users of the instrument to target groups of students in need of support with respect to one factor but not the other, recognising that practical strategies used to seek feedback are different to those used to enact it.

The results of the Rasch analysis were satisfactory: there were no significantly misfitting items in the final scale, no disordered thresholds in the six-point rating scale, and no substantial DIF across gender. The detailed Rasch analysis results are placed in supplementary online material.

The Cronbach's alpha coefficients for subscales SF, MS, UF, PF and MA were 0.66, 0.64, 0.76, 0.69 and 0.81 respectively. As Cronbach's alpha represents a lower bound to the reliability and in many cases underestimates the true reliability (Sijtsma 2009), these results appear acceptable, although there is room for further improvement. The omega coefficients for the five subscales were 0.67, 0.64, 0.76, 0.69 and 0.81, which were quite similar to Cronbach's alpha coefficients.

When a subsample of the participants ($N = 322$, 92%) completed the FLBS four weeks after the main survey, the Cronbach's alphas and omega coefficients for subscale scores were similar (see Table 2), indicating the measurement error is acceptable (Nunnally and Bernstein 1994). The test–retest correlations of participants' responses for all subscales were positive and significant (ranging from 0.56 to 0.71), indicating fair to good test-retest reliability according to Cicchetti's (1994) guideline. The moderate stability of students' feedback literacy over time was in line with our expectations because, on the one hand, feedback literacy represents a long-term capacity which may remain stable and, on the other hand, the FLBS focused on the behaviour elements of feedback literacy that are susceptive to change due to learning, practice and context.

To test the external validity of the FLBS, Pearson correlations between the Rasch-calibrated person measures of the five FLBS subscales and the five BFI-S subscales were calculated; these are reported in Table 3.

The Cronbach's alphas of the five BFI-S subscales – extraversion, agreeableness, conscientiousness, emotional stability and openness – were 0.85, 0.78, 0.71, 0.77 and 0.70 respectively.

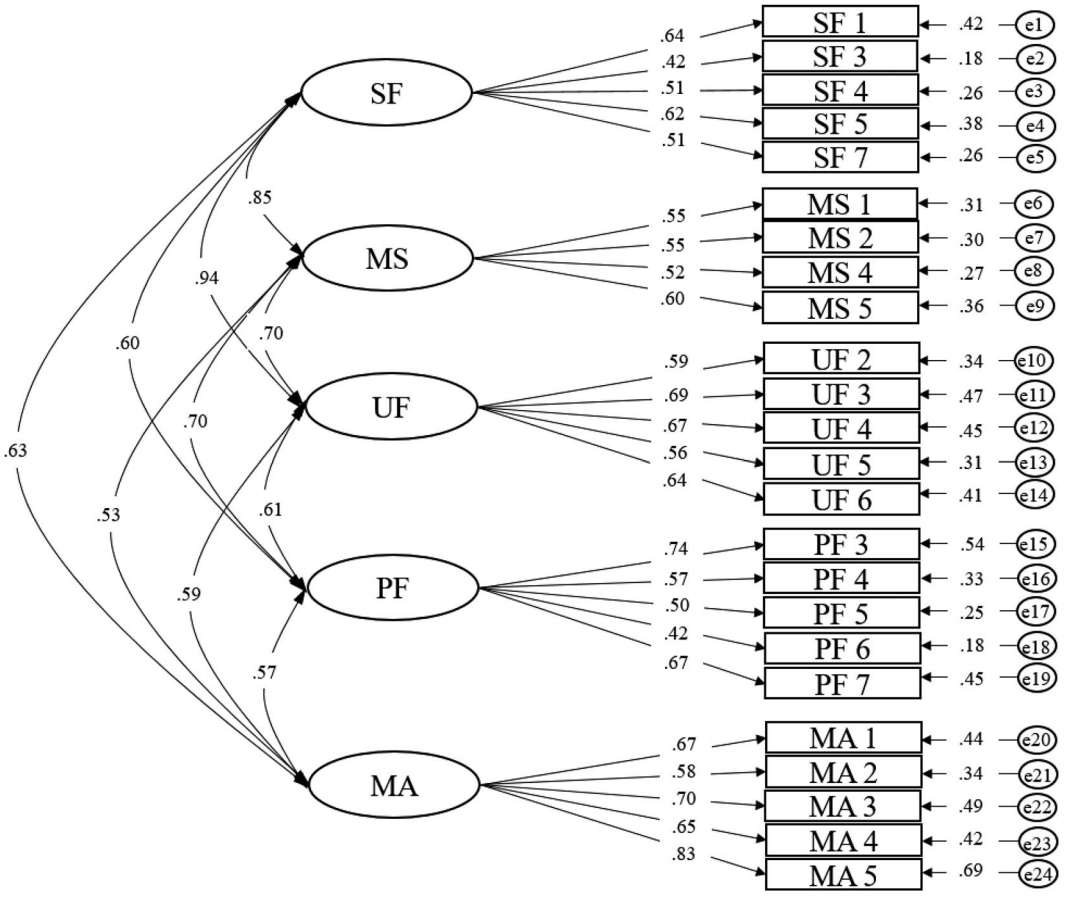

Figure 2. The five-factor **FLBS** CFA model.
SF: Seek Feedback information, MS: Make Sense of information, UF: Use Feedback information, PF: Provide Feedback information, MA: Manage Affect.

Table 2. Cronbach's alphas for subscale scores and test-retest correlations for subsample ($N = 322$).

	T1		T2		
Subscale	α	ω	α	ω	Test–retest correlations
SF	0.66	0.67	0.68	0.68	0.61**
MS	0.64	0.64	0.70	0.71	0.59**
UF	0.76	0.76	0.74	0.74	0.58**
PF	0.69	0.69	0.65	0.65	0.60**
MA	0.81	0.81	0.79	0.79	0.66**

**$p < 0.01$

Discussion

Our results provide initial validity evidence for the Feedback Literacy Behaviour Scale as a measure of students' self-reported enactment of their feedback literacy. The five-factor model was generally supported by both the factor analysis and Rasch analysis, and the instrument has an acceptable level of test-retest reliability. This represents an important first step into more robust quantitative research on feedback literacy behaviours. Researchers, educators and learners wishing to measure feedback literacy behaviours can access the instrument in Appendix A of this paper.

Table 3. Pearson correlations between FLBS and BFI-S subscales.

	SF	MS	UF	PF	MA
Extraversion	.19**	.09	.15**	.22**	.10
Agreeableness	.15**	.10	.23**	.22**	.00
Conscientiousness	.16**	.06	.18**	.18**	.14**
Emotional stability	−0.01	.08	.09	.02	.31**
Openness	.14**	.16**	.08	.18**	.14*

*$p < 0.05$.
**$p < 0.01$.

One issue requiring further exploration is whether the capability to provide feedback is a necessary component of feedback literacy. Our conceptual framework, informed by Carless and Boud (2018) and Molloy, Boud, and Henderson (2020) suggests that, at its core, feedback is a reciprocal act where learners are active participants who also generate feedback information. In such an understanding of feedback, being feedback literate requires the capability to provide information to others. However, given our expert review of the initial scale suggested that our experts did not all agree, and the inconsistency among existing scales on this matter (Liao 2021; Zhan 2022; Song 2022; Yildiz, Bozpolat, and Hazar 2022; Yu, Di Zhang, and Liu 2022; Dong, Gao, and Schunn 2023), there is a need for further work in this space. Could it be that this is a value position that assumes that learners must be equipped to help others as well as themselves, or is it a matter that can be separated in more teacher-centred pedagogic contexts? The reasonably high correlations between *provide feedback* (PF) and other factors provides some initial evidence to support our view that feedback literacy incorporates the provision of feedback information.

Another conceptual matter raised by our development and validation process is the close correlation between *seek feedback* (SF) and *use feedback* (UF). We opted to keep these factors separate for reasons discussed earlier, but their closeness suggests something that is perhaps obvious but still conceptually useful: those students who seek feedback were likely to also use feedback. However, these factors remain conceptually distinct in our model, supported by the evidence of patterns of correlations among the FLBS subscales and the Big Five dimensions.

Overall, links between personality and the FLBS were of the expected direction (Lipnevich et al. 2021), thus providing additional validity evidence and suggesting the construct's differentiation from the Big Five personality factors. Conscientiousness and openness were the strongest predictors of behavioural indicators of students' engagement with feedback, revealing significant correlations with four out of five FLBS subscales. It suggests that students who tend to be disciplined and achievement-focused (high on conscientiousness) as well as open to new information (high on openness) would tend to be higher on feedback literacy. Agreeableness yielded significant links with SF, UF and PF, showing that students who were more cooperative and trusting were more likely to seek, use and provide feedback. Out of five FLBS factors, emotional stability was significantly and positively related to students' ability to manage their affect (MA). This correlation is of the highest magnitude among all observed links ($r = 0.311$, $p < 0.01$), yet it is weak enough to show evidence of factor differentiation.

Of particular interest are the links between our highly correlated factors of SF and UF and the Big Five factors. They show that students' willingness to use feedback (UF) has a stronger relation with their tendency to be trusting (agreeableness) compared to their tendency to be disciplined (conscientiousness). The reverse is true for feedback seeking (SF): there was a higher correlation between conscientiousness and SF than between agreeableness and SF. Feedback use (UF) was the only factor unrelated to openness, showing that one's proclivity to be intellectually curious doesn't matter as much in one's use of feedback; this is not the case for feedback seeking (SF), when intellectual curiosity apparently matters. These differential relations support our separation of the UF and SF factors, although further research is needed.

In addition to its focus on behaviours, the FLBS differs from existing instruments (Liao 2021; Zhan 2022; Song 2022; Yildiz, Bozpolat, and Hazar 2022; Yu, Di Zhang, and Liu 2022; Dong, Gao,

and Schunn 2023) in that it does not use language specific to the context of education. This means the FLBS may have utility beyond education, and it may enable longitudinal tracking of feedback literacy behaviours from late high school, through university and onto graduate employment settings. However, this would require further validation of the instrument in those contexts.

The literature on feedback literacy interventions is expanding (Little et al. 2023), including approaches such as self and peer assessment (Hoo, Deneen, and Boud 2022), mixtures of online and face-to-face workshops (Noble et al. 2019), and freely available toolkits (Winstone, Mathlin, and Nash 2019). Previous work reviewing students' proactive recipience of feedback found that meta-analysis was not possible, partly due to methodological inconsistency (Winstone, Nash, Parker, et al. 2017). The FLBS may help avoid this problem in future feedback literacy research by providing a common measure specifically focused on behaviour. The FLBS could facilitate not only studies that more reliably measure students' feedback literacy behaviours before and after an intervention, but also comparison of those results with other studies that use the FLBS. It may also prove a useful tool for longitudinal studies of feedback literacy behaviours.

Researchers and educators may wish to consider the five-factor model of feedback literacy underpinning the FLBS when designing interventions to improve feedback literacy behaviours. It may be that targeted development of particular aspects of feedback literacy is needed. Administering the FLBS would help identify cohorts that need help with, for example, managing affect, but not with making sense of feedback.

Limitations and future directions

There are potential limitations to the use of our instrument. The first is that it relies on students to self-report what they do. Further mixed-methods work is required to understand how accurate these self-reports are. As our scale focuses on feedback literacy behaviours, such work would include observations of student feedback processes. One weakness regarding the psychometric properties of the scale is the relatively low Cronbach's alphas and omega coefficients for subscales SF and MS. If this level of reliability was duplicated on other samples, future studies may consider further developing the scale – such as adding additional items representing other relevant feedback practices to these two subscales – with an aim of increased reliability. Another major challenge is that the degree to which feedback literacy is context-dependent is yet to be established, so we do not know how transferrable the results of the FLBS are across contexts. Future work might involve validating the instrument within different disciplines, age groups and countries, and how well the self-reported behaviours in the instrument map to observable behaviours in practice.

Conclusion

Feedback literacy helps students engage with the challenges and opportunities that feedback offers. The Feedback Literacy Behaviour Scale (FLBS) provides a way to capture the current state of students' feedback literacy behaviours, and to measure changes over time. It is intended to be useful for researchers seeking to study the effects of interventions to improve feedback literacy, as well as for everyday educators and students. As a measure of feedback literacy behaviours rather than attitudes or conceptions, the FLBS is targeted at what students do in feedback processes – which, according to recent understandings of feedback, is what matters most.

Acknowledgements

We wish to thank Kieran Balloo, Margaret Bearman, Jaclyn Broadbent, Sin Wang Chong, Michael Henderson, Gordon Joughin, Rob Nash, Karen Olave-Encina, Edd Pitt, Bas Trimbos, Lan Yang and the other experts who contributed to the expert review of our initial scale.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Phillip Dawson http://orcid.org/0000-0002-4513-8287
Zi Yan http://orcid.org/0000-0001-9305-884X
Anastasiya Lipnevich http://orcid.org/0000-0003-0190-8689
Joanna Tai http://orcid.org/0000-0002-8984-2671
David Boud http://orcid.org/0000-0002-6883-2722
Paige Mahoney http://orcid.org/0000-0003-4719-6015

References

Andrich, D. 2004. "Controversy and the Rasch Model: A Characteristic of Incompatible Paradigms?" *Medical Care* 42 (1 Suppl): I7–16. doi:10.1097/01.mlr.0000103528.48582.7c.

Anseel, F., A. S. Beatty, W. Shen, F. Lievens, and P. R. Sackett. 2015. "How Are We Doing after 30 Years? A Meta-Analytic Review of the Antecedents and Outcomes of Feedback-Seeking Behavior." *Journal of Management* 41 (1): 318–348. doi:10.1177/0149206313484521.

Arbuckle, J. L. 2015. *Amos (Version 24.0)[Computer Program]*. Chicago: IBM SpSS.

Bond, T., Z. Yan, and M. Heene. 2020. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. New York: Routledge.

Boud, D., and E. Molloy. 2013. "What is the Problem with Feedback?." In *Feedback in Higher and Professional Education: Understanding It and Doing It Well*, edited by D. Boud and E. Molloy, 1–10. Abingdon: Routledge.

Briley, D. A., M. Domiteaux, and E. M. Tucker-Drob. 2014. "Achievement-Relevant Personality: Relations with the Big Five and Validation of an Efficient Instrument." *Learning and Individual Differences* 32: 26–39. doi:10.1016/j.lindif.2014.03.010.

Brown, G. T. L., E. R. Peterson, and E. S. Yao. 2016. "Student Conceptions of Feedback: Impact on Self-Regulation, Self-Efficacy, and Academic Achievement." *The British Journal of Educational Psychology* 86 (4): 606–629. doi:10.1111/bjep.12126.

Carless, D., and D. Boud. 2018. "The Development of Student Feedback Literacy: Enabling Uptake of Feedback." *Assessment & Evaluation in Higher Education* 43 (8): 1315–1325. doi:10.1080/02602938.2018.1463354.

Chong, S. W. 2021. "Reconsidering Student Feedback Literacy from an Ecological Perspective." *Assessment & Evaluation in Higher Education* 46 (1): 92–104. doi:10.1080/02602938.2020.1730765.

Cicchetti, D. V. 1994. "Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology." *Psychological Assessment* 6 (4): 284–290. doi:10.1037/1040-3590.6.4.284.

Cronbach, L. J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16 (3): 297–334. doi:10.1007/BF02310555.

Deneen, C., G. T. L. Brown, T. G. Bond, and R. Shroff. 2013. "Understanding Outcome-Based Education Changes in Teacher Education: Evaluation of a New Instrument with Preliminary Findings." *Asia-Pacific Journal of Teacher Education* 41 (4): 441–456. doi:10.1080/1359866X.2013.787392.

Dong, Z., Y. Gao, and C. D. Schunn. 2023. "Assessing Students' Peer Feedback Literacy in Writing: Scale Development and Validation." *Assessment & Evaluation in Higher Education*. doi:10.1080/02602938.2023.2175781.

Donnellan, M. B., F. L. Oswald, B. M. Baird, and R. E. Lucas. 2006. "The mini-IPIP Scales: Tiny-yet-Effective Measures of the Big Five Factors of Personality." *Psychological Assessment* 18 (2): 192–203. doi:10.1037/1040-3590.18.2.192.

Goldberg, L. R. 1993. "The Structure of Phenotypic Personality Traits." *The American Psychologist* 48 (1): 26–34. doi:10.1037//0003-066x.48.1.26.

Gorsuch, R. L. 1983. *Factor Analysis*. 2nd ed. Hillsdale, NJ: Erlbaum Associates.

Gravett, K. 2022. "Feedback Literacies as Sociomaterial Practice." *Critical Studies in Education* 63 (2): 261–274. doi:10.1080/17508487.2020.1747099.

Hattie, J., and H. Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112. doi:10.3102/003465430298487.

Hinkin, T. R. 1998. "A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires." *Organizational Research Methods* 1 (1): 104–121. doi:10.1177/109442819800100106.

Hoo, H.-T., C. Deneen, and D. Boud. 2022. "Developing Student Feedback Literacy through Self and Peer Assessment Interventions." *Assessment & Evaluation in Higher Education* 47 (3): 444–457. doi:10.1080/02602938.2021.1925871.

Hu, L. t., and P. M. Bentler. 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. doi:10.1080/10705519909540118.

Johns, R. 2005. "One Size Doesn't Fit All: Selecting Response Scales for Attitude Items." *Journal of Elections, Public Opinion & Parties* 15 (2): 237–264. doi:10.1080/13689880500178849.

Joughin, G., D. Boud, P. Dawson, and J. Tai. 2021. "What Can Higher Education Learn from Feedback Seeking Behaviour in Organisations? Implications for Feedback Literacy." *Assessment & Evaluation in Higher Education* 46 (1): 80–91. doi:10.1080/02602938.2020.1733491.

Krasman, J. 2010. "The Feedback-Seeking Personality: Big Five and Feedback-Seeking Behavior." *Journal of Leadership & Organizational Studies* 17 (1): 18–32. doi:10.1177/1548051809350895.

Liao, Y. 2021. "Developing and Validating the Student Assessment-Based Feedback Literacy (SAFL) Instrument: A Mixed-Methods Approach." PhD, University of Texas at El Paso.

Linderbaum, B. A., and P. E. Levy. 2010. "The Development and Validation of the Feedback Orientation Scale (FOS)." *Journal of Management* 36 (6): 1372–1405. doi:10.1177/0149206310373145.

Lipnevich, A. A., K. Gjicali, M. Asil, and J. K. Smith. 2021. "Development of a Measure of Receptivity to Instructional Feedback and Examination of Its Links to Personality." *Personality and Individual Differences* 169: 110086. doi:10.1016/j.paid.2020.110086.

Lipnevich, A. A., and C. Lopera-Oquendo. 2022. "Receptivity to Instructional Feedback: A Validation Study in the Secondary School Context in Singapore." *European Journal of Psychological Assessment*.

Little, T., P. Dawson, D. Boud, and J. Tai. 2023. "Can Students' Feedback Literacy Be Improved? A Scoping Review of Interventions." *Assessment & Evaluation in Higher Education*. doi:10.1080/02602938.2023.2177613.

MacCann, C., A. A. Lipnevich, J. Burrus, and R. D. Roberts. 2012. "The Best Years of Our Lives? Coping with Stress Predicts School Grades, Life Satisfaction, and Feelings about High School." *Learning and Individual Differences* 22 (2): 235–241. doi:10.1016/j.lindif.2011.08.004.

McDonald, R. P. 1999. *Test Theory: A Unified Treatment*. Mahwah, NJ: Taylor & Francis.

McDonald, R. P., and M.-H. R. Ho. 2002. "Principles and Practice in Reporting Structural Equation Analyses." *Psychological Methods* 7 (1): 64–82. doi:10.1037/1082-989x.7.1.64.

Messick, S. 1995. "Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning." *American Psychologist* 50 (9): 741–749. doi:10.1037/0003-066X.50.9.741.

Molloy, E., D. Boud, and M. Henderson. 2020. "Developing a Learning-Centred Framework for Feedback Literacy." *Assessment & Evaluation in Higher Education* 45 (4): 527–540. doi:10.1080/02602938.2019.1667955.

Moors, G. 2008. "Exploring the Effect of a Middle Response Category on Response Style in Attitude Measurement." *Quality & Quantity* 42 (6): 779–794. doi:10.1007/s11135-006-9067-x.

Mueller, R. O., and G. R. Hancock. 2001. "Factor Analysis and Latent Structure: Confirmatory Factor Analysis." In *International Encyclopedia of the Social and Behavioral Sciences*, edited by N. J. Smelser and P. B. Baltes, 5239–5244. Oxford, UK: Elsevier.

Noble, C., C. Sly, L. Collier, L. Armit, J. Hilder, and E. Molloy. 2019. "Enhancing Feedback Literacy in the Workplace: A Learner-Centred Approach." In *Augmenting Health and Social Care Students' Clinical Learning Experiences: Outcomes and Processes*, edited by S. Billett, J. Newton, G. Rogers, and C. Noble, 283–306. Cham: Springer International Publishing.

Nunnally, J. C., and I. H. Bernstein. 1994. *Psychometric Theory*. New York, NY: McGraw-Hill.

Pintrich, P. R., D. A. F. Smith, T. Garcia, and W. J. McKeachie. 1993. "Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (Mslq)." *Educational and Psychological Measurement* 53 (3): 801–813. doi:10.1177/0013164493053003024.

Rasch, G. 1960. Studies in Mathematical Psychology: I Probabilistic Models for Some Intelligence and Attainment Tests. Expanded ed. 1980. Chicago, IL: The University of Chicago Press.

Saris, W. E., A. Satorra, and W. M. Van der Veld. 2009. "Testing Structural Equation Models or Detection of Misspecifications?" *Structural Equation Modeling: A Multidisciplinary Journal* 16 (4): 561–582. doi:10.1080/10705510903203433.

Sijtsma, K. 2009. "On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha." *Psychometrika* 74 (1): 107–120. doi:10.1007/s11336-008-9101-0.

Song, B. K. 2022. "Bifactor Modelling of the Psychological Constructs of Learner Feedback Literacy: Conceptions of Feedback, Feedback Trust and Self-Efficacy." *Assessment & Evaluation in Higher Education* 47 (8): 1444–1457. doi:10.1080/02602938.2022.2042187.

Testa, I., G. Capasso, A. Colantonio, S. Galano, I. Marzoli, U. Scotti di Uccio, F. Trani, and A. Zappia. 2019. "Development and Validation of a University Students' Progression in Learning Quantum Mechanics through Exploratory Factor Analysis and Rasch Analysis." *International Journal of Science Education* 41 (3): 388–417. doi:10.1080/09500693.2018.1556414.

Trapmann, S., B. Hell, J.-O W. Hirn, and H. Schuler. 2007. "Meta-Analysis of the Relationship between the Big Five and Academic Success at University." *Zeitschrift für Psychologie/Journal of Psychology* 215 (2): 132–151. doi:10.1027/0044-3409.215.2.132.

VandeWalle, D., S. Ganesan, G. N. Challagalla, and S. P. Brown. 2000. "An Integrated Model of Feedback-Seeking Behavior: Disposition, Context, and Cognition." *The Journal of Applied Psychology* 85 (6): 996–1003. doi:10.1037/0021-9010.85.6.996.

Walton, K. E., D. Murano, J. Burrus, and A. Casillas. 2023. "Multimethod Support for Using the Big Five Framework to Organize Social and Emotional Skills." *Assessment* 30 (1): 144–159. doi:10.1177/10731911211045744.
Winstone, N. E., R. Ajjawi, K. Dirkx, and D. Boud. 2022. "Measuring What Matters: The Positioning of Students in Feedback Processes within National Student Satisfaction Surveys." *Studies in Higher Education* 47 (7): 1524–1536. doi:10.1080/03075079.2021.1916909.
Winstone, N. E., D. Boud, P. Dawson, and M. Heron. 2022. "From Feedback-as-Information to Feedback-as-Process: A Linguistic Analysis of the Feedback Literature." *Assessment & Evaluation in Higher Education* 47 (2): 213–230. doi:10.1080/02602938.2021.1902467.
Winstone, N. E., and D. Carless. 2019. *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. London: Routledge.
Winstone, N. E., G. Mathlin, and R. A. Nash. 2019. "Building Feedback Literacy: Students' Perceptions of the Developing Engagement with Feedback Toolkit." *Frontiers in Education* 4 (39): 1–11. doi:10.3389/feduc.2019.00039.
Winstone, N. E., R. A. Nash, M. Parker, and J. Rowntree. 2017. "Supporting Learners' Agentic Engagement with Feedback: A Systematic Review and a Taxonomy of Recipience Processes." *Educational Psychologist* 52 (1): 17–37. doi:10.1080/00461520.2016.1207538.
Winstone, N. E., R. A. Nash, J. Rowntree, and M. Parker. 2017. "It'd Be Useful, but I Wouldn't Use It': barriers to University Students' Feedback Seeking and Recipience." *Studies in Higher Education* 42 (11): 2026–2041. doi:10.1080/03075079.2015.1130032.
Wisniewski, B., K. Zierer, and J. Hattie. 2019. "The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research." *Frontiers in Psychology* 10: 3087. doi:10.3389/fpsyg.2019.03087.
Yan, Z., S. Brubacher, D. Boud, and M. Powell. 2020. "Psychometric Properties of the Self-Assessment Practice Scale for Professional Training Contexts: Evidence from Confirmatory Factor Analysis and Rasch Analysis." *International Journal of Training and Development* 24 (4): 357–373. doi:10.1111/ijtd.12201.
Yildiz, H., E. Bozpolat, and E. Hazar. 2022. "Feedback Literacy Scale: A Study of Validation and Reliability." *International Journal of Eurasian Education and Culture* 7 (19): 2214–2249.
Yu, S., E. Di Zhang, and C. Liu. 2022. "Assessing L2 Student Writing Feedback Literacy: A Scale Development and Validation Study." *Assessing Writing* 53: 100643. doi:10.1016/j.asw.2022.100643.
Zhan, Y. 2022. "Developing and Validating a Student Feedback Literacy Scale." *Assessment & Evaluation in Higher Education* 47 (7): 1087–1100. doi:10.1080/02602938.2021.2001430.

Appendix A. The feedback literacy scale (FLBS)

Feedback Literacy Behaviour Scale (final version)

For the next series of questions, please think about what you usually do in your studies, and rate how often you do these things.

Seek feedback information (SF)

SF1: I reflect on the quality of my own work and use my reflection as a source of information to improve my work.
SF3: I seek out examples of good work to improve my work.
SF4: When other people provide me with input about my work I listen or read thoughtfully.
SF5: When I am working on a task, I consider comments I have received on similar tasks.
SF7: I ask for comments about specific aspects of my work.

Make sense of information (MS)

MS1: I carefully consider comments about my work before deciding if I will use them or not.
MS2: When receiving conflicting information about my work from different sources, I make a judgement about what I will use.
MS4: When deciding what to do with comments, I consider the credibility of their sources.
MS5: I consider how comments relate to criteria or standards.

Use feedback information (UF)

UF2: I check whether my work is better after I have acted on comments.
UF3: I use comments on my work to refine my understanding of what good quality work looks like.
UF4: When receiving comments I plan how I will use them to improve my future work, not just the immediate task.
UF5: I keep comments on previous work to use again in the future.
UF6: When I receive comments from others, I use them to improve what I'm working on at the time.

Provide feedback information (PF)

PF3: When commenting on the work of others, I provide constructive criticism.
PF4: I comment on other people's work when I am invited to.
PF5: When commenting on other people's work I refer to standards or criteria.
PF6: I offer to provide feedback to my peers.

PF7: I try to be very clear when providing feedback comments to others.

Manage affect (MA)

MA1: I am open to reasonable criticism about my work.

MA2: I deal well with any negative emotional responses I have to feedback information.

MA3: When a feedback message is valuable but upsetting or annoying, I still find a way to make use of it.

MA4: Critical comments motivate me to improve my work.

MA5: I make use of critical comments even if they are difficult to receive.