Student perceptions of assessment feedback: a critical scoping review and call for research

Fabienne M. Van der Kleij & Anastasiya A. Lipnevich

Educational Assessment, Evaluation and Accountability
International Journal of Policy, Practice and Research

ISSN 1874-8597

Educ Asse Eval Acc
DOI 10.1007/s11092-020-09331-x

Springer

Educational Assessment, Evaluation and Accountability
https://doi.org/10.1007/s11092-020-09331-x

Student perceptions of assessment feedback: a critical scoping review and call for research

Fabienne M. Van der Kleij[1,2] · Anastasiya A. Lipnevich[3]

Received: 7 April 2020 / Accepted: 29 July 2020/Published online: 17 August 2020

Abstract

The potential of feedback to enhance students' performance on a task, strategies, or learning has long been recognized in the literature. However, feedback needs to be utilized by a learner to realize its potential. Hence, examining student perceptions of feedback and their links to effective uptake of feedback has been the focus of much recent feedback research. This paper presents a critical scoping review of the feedback perceptions literature. The review discusses the methods employed by 164 studies published between 1987 and 2018 and synthesizes the main findings across this body of literature. Lacking theoretical frameworks, repetitiveness (not replicability) of studies, and methodological problems observed among the reviewed have resulted in somewhat disappointing conclusions. Based on the findings, we present a framework for future investigations into student perceptions of feedback and suggest several avenues for the future of the field.

Keywords Feedback · Student perceptions · Performance · Scoping review · Methodology

1 Introduction

The key role of feedback in daily instructional practices has been studied extensively over the past few decades. Handbooks, meta-analyses, and compendiums of reviewed literature have come to a consistent conclusion: feedback is critical to student learning

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11092-020-09331-x) contains supplementary material, which is available to authorized users.

✉ Fabienne M. Van der Kleij
f.vanderkleij@uq.edu.au

[1] Institute for Learning Sciences & Teacher Education, Australian Catholic University, Level 4, 229 Elizabeth Street, Brisbane, Queensland 4000, Australia

[2] Present address: School of Education, The University of Queensland, St Lucia, Queensland 4072, Australia

[3] Queens College and the Graduate Center, City University of New York, New York, USA

and performance. Hattie, for example, identified feedback as one of the most powerful influences on student learning outcomes (Hattie, 2009). However, there are specific conditions, factors, and processes that make feedback more or less conducive to learning improvements. Such factors often get overlooked in the general discussion of feedback effectiveness.

Student perceptions of feedback have been identified as an important factor in determining the impact of feedback on learning (e.g. Carless, 2006; Havnes, Smith, Dysthe, & Ludvigsen, 2012; Jonsson, 2013; Lipnevich & Smith, 2009; Rakoczy, Harks, Klieme, Blum, & Hochweber, 2013; Wellington, 2010). If a student does not perceive feedback as useful, helpful, objective, or relevant, there is hardly a chance that this feedback will have any bearing on students' subsequent work (Jonsson & Panadero, 2018; Lipnevich & Smith, 2009). In recent years, much research has emerged focusing on student feedback perceptions. In this critical scoping review, we synthesized literature on student perceptions of feedback with the goal to offer suggestions for future investigations, by (1) describing and discussing methods used to investigate student perceptions of feedback and bring attention to methodological strengths and limitations of these methods and (2) summarizing main findings from this area of research to identify potential gaps in our understanding of student perceptions of feedback.

1.1 Defining feedback

Definitions of feedback in educational contexts have evolved considerably over the past century (e.g. Lipnevich & Smith, 2018; Mory, 1996; Van der Kleij, Adie, & Cumming, 2019). Early definitions in a behaviourist paradigm viewed feedback as a reinforcing message, which provided the critical link between stimuli and correct responses. Students were considered as playing a passive role, and the outcomes of providing feedback were considered to be predictable. From the 1970s, the information processing view of feedback gained popularity. Within this perspective, feedback was construed as information that learners could use to correct errors (Mory, 1996), thus acknowledging that students play a role in deciding (how) to use feedback information.

A bulk of current feedback research still refers to Ramaprasad's (1983) definition of feedback, in which he described feedback as information about where (a) a student is, where (b) the student is going, and how to get from point *a* to *b*. This definition was conceptualized in an educational context firstly by Sadler (1989), and later by Hattie and Timperley (2007). Their conceptualizations position the student as a critical player in the feedback process. That is, these conceptualizations consider that if feedback is to affect student learning, the learner needs to actively use the feedback information to make decisions and take appropriate action. However, the extent to which the student plays an active role in the feedback process varies even within conceptualizations of feedback that consider the critical role of the student (Van der Kleij et al., 2019). Some researchers have interpreted this to mean that students need to be told how to improve and take corrective action, which is 'a very linear and procedural approach to feedback' (Torrance, 2012, p. 333). Others have taken a more student-centred approach, considering that feedback involves 'dialogic processes whereby learners make sense of information from various sources and use it to enhance their work or learning strategies' (Carless, 2016, p. 1).

Consistent with these widely used conceptualizations of feedback, in this paper, we follow the definition provided by Smith and Lipnevich (2018), who defined feedback as any information about a performance that learners can use to improve their performance or learning. Feedback might come from teachers, peers, or the task itself; it can be computer or self-generated by the learner. It may include information on where the learner is, where the learner is going, or what steps should be taken and strategies employed to get there. This definition applies to both immediate and specific effects of feedback, as well as to more general and long-term influences of feedback (Lipnevich, Berg, & Smith, 2016). Consistent with the writings of Ramaprasad (1983), Sadler (1989), and Hattie and Timperley (2007), it also brings the learner into the focus and emphasizes the fact that feedback needs to be utilized by a learner to enhance his or her performance on a task, strategies, or learning. Hence, examining student perceptions of feedback and their links to effective uptake of feedback is in order, which has been the focus of much recent feedback research.

1.2 Feedback effectiveness and formative assessment

Before we proceed with close examination of student perceptions of feedback, we would like to situate feedback within the realm of formative assessment. Guskey (2018) reminds us that it was Bloom who took Scriven's distinction of summative and formative evaluation and applied it to the context of classroom assessment (Bloom, 1968). Bloom suggested that feedback was an indelible part of formative assessment and the purpose of it was to improve student learning. This idea is a leitmotif that is woven into all reviews of feedback (e.g. Hattie & Timperley, 2007; Lipnevich et al., 2016; Shute, 2008; Smith & Lipnevich, 2018). In other words, formative assessment is one of the approaches through which feedback is provided to the learner and its sole goal is to help students improve their performance on a task and/or learning.

There are plentiful research studies examining the effectiveness of feedback for students' academic attainment (e.g. reviews by Black & Wiliam, 1998; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Shute, 2008; Lipnevich & Smith, 2018). Such research has demonstrated that in general, certain types of feedback were associated with enhanced learning outcomes. For instance, providing explanations or identifying areas for improvement following assessment is much more powerful than providing simple corrective information, numerical grades, or praise (Hattie & Timperley, 2007; Lipnevich & Smith, 2009; Shute, 2008; Van der Kleij, Feskens, & Eggen, 2015).

However, the relation between feedback and learning is still poorly understood, and past research sheds limited light on the specific mechanisms of feedback and its links to meaningful educational outcomes. As numerous reviews of feedback effectiveness demonstrate, positive effects of feedback on learning cannot be guaranteed (e.g. Kluger & DeNisi, 1996; Shute, 2008; Van der Kleij et al., 2015). More specifically, and emphasizing the aforementioned point about the key role of a student, no matter how perfect the feedback message is, if a student is not able or willing to use it, one cannot anticipate any improvement. In the 1970s, following decades of inconsistent research findings on feedback effects on learning, the field gradually came to a realization that students must somehow 'process' feedback information mindfully (see reviews by Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Mory, 1996) in order for it to impact

their learning. Several conceptual models have been proposed to understand *how* feedback influences learning (see e.g. Bangert-Drowns et al., 1991; Kluger & DeNisi, 1996). However, the focus of much of the feedback research remained on identifying how to provide 'effective' feedback (Shute, 2008), with limited consideration of the student perspective, and their agency in deciding whether and how to *engage* with feedback (Handley, Price, & Millar, 2011; Lipnevich et al., 2016).

Recent education research increasingly recognizes that students' engagement with instructional feedback is a necessary prerequisite for feedback to positively affect learning (e.g. Handley et al., 2011; Lipnevich et al., 2016; Winstone, Nash, Parker, & Rowntree, 2017). Black and Wiliam (1998) warned that:

> in relation to action taken by the learner it would be a mistake to regard the student as the passive recipient of a call to action. There are complex links between the way in which the message is received, the way in which that perception motivates a selection amongst different courses of action, and the learning activity which may or may not follow. (pp. 20–21)

In other words, feedback only has the potential to enhance student learning if it makes students realize they need to adapt their beliefs, knowledge, or strategies and take appropriate actions based on the feedback (Black & Wiliam, 2009; Butler & Winne, 1995), that is, engage with the feedback. Thus, the way students perceive feedback determines their affective, cognitive, and behavioural responses to it.

Lipnevich et al. (2016) proposed a feedback–student interaction model that may be useful in considering the complexity of feedback and factors that may affect student perceptions and subsequent action (or lack thereof). According to their model, feedback is always received in a context that may be more or less familiar and comfortable and where the outcomes would be perceived as important or less so. Hence, the consequences of such feedback will be variable. Further, feedback may vary in tone, length, specificity, and complexity (among other characteristics) and, consequently, elicit differential perceptions, appraisals, emotions, and behavioural responses (Goetz, Lipnevich, Krannich, & Gogol, 2018). For example, supportive feedback would have a higher chance of eliciting enjoyment and pride, whereas feedback evoking social comparisons and feedback that is judgmental in tone will have a higher chance of eliciting anxiety. Students may be joyful or sad, confused or inspired, and, as a result, they may act adaptively, maladaptively, or not respond at all. Feedback may trigger responses that are very specific to the task or that are generalizable across domains. These complex contingencies among feedback, context, and the student significantly influence whether or not the message that a student receives and considers will be actionable and whether it will be used to improve performance and learning. According to this model, student perception of feedback is the first step in the cycle of effective engagement with feedback.

1.3 Perceptions of feedback as a precursor to engagement

Engagement with feedback includes students' perception, interpretation, and the actual use of feedback (Handley et al., 2011; Hargreaves, 2011; Nicol, 2013). Handley et al. (2011), for example, distinguish between (1) 'readiness-to-engage' (p. 550) with

feedback, a mindset prior to receiving feedback that is influenced by performance expectations and prior feedback experiences, and 'active engagement' (p. 549) that includes both visible and invisible thoughts and actions following feedback. Based on the literature published thus far, a preliminary understanding of what needs to be in place for students to successfully engage with feedback to improve their learning involves at least the following aspects (Carless, 2006; Gamlem & Smith, 2013; Handley et al., 2011; Hargreaves, 2011; Havnes et al., 2012; Jonsson & Panadero, 2018; Nicol, 2013): willingness to attend to it, noticing feedback, understanding feedback (or asking for clarification), perceiving the feedback as relevant and useful, reflecting on feedback in relation to their own work, having time and opportunity to use it, and being able to translate it into a plan for action (independently or with help from the teacher) or use it to directly to revise knowledge networks.

Jonsson (2013) reviewed the literature on students' *use* of feedback in the context of higher education. Although not explicitly conceptualized as such, the term 'feedback use' seemed to encompass the broader aspects of engaging with feedback. In his review, only two studies were identified that investigated students' *actual use* of feedback. The other 101 reviewed studies employed post hoc strategies, such as analysis of work samples and revisions, and reflective interviews. From this indirect evidence of feedback use, Jonsson (2013) identified five key reasons for the lack of feedback use among students: (a) it is not (perceived as) useful, (b) it contains insufficient detail or is not sufficiently individualized, (c) it is too authoritative in tone, (d) students lack the ability to effectively use it, and (e) it is provided in terminology that students do not understand. These conditions correspond to the aspects required for student engagement with feedback listed above. In 2017, Winstone and colleagues reviewed factors that may affect student use of feedback and discussed moderators of student engagement with feedback. The latter included perceived credibility of the feedback provider, the mode, timing, and content of the feedback message, student characteristics, as well as contextual factors. The researchers discussed factors, which, alone and in combination, explain how feedback is perceived and subsequently acted upon by a student.

Jonsson and Panadero (2018), in their review of factors contributing to student active engagement with feedback, list the following three conditions as critical to learners' productive use of feedback: (1) feedback is perceived as useful, (2) students possess strategies to utilize feedback, and (3) feedback does not serve as a basis for social comparison (i.e. excludes grades). The first condition is of particular interest for our review. We would like to emphasize that the current review focuses on student *perceptions* of feedback rather than their *conceptions* thereof. The term *feedback perception* captures how students comprehend, perceive, and value a feedback message and how they experience and receive feedback both from cognitive and affective perspectives, as opposed to students' conceptualizations of the meaning of feedback more generally (i.e. conceptions; see Brown, 2011).

1.4 Methodological issues

Despite the fact that a significant number of research investigations into student perceptions of feedback has emerged in recent years, particularly in higher education contexts, our understanding of this issue is still incomplete. This may be attributed to

particular methods employed in studies investigating student feedback perceptions, which do not allow for causal claims. More specifically, our understanding of student perceptions and their potential contribution to students' effective uptake of feedback as well as possible effects on performance remains limited. This is a critical issue, as many of these studies are founded upon the idea that positive student perceptions are associated with effective uptake of feedback and vice versa, as suggested by, among others, Jonsson (2013) and Jonsson and Panadero (2018).

Even more importantly, previous reviews (e.g. Jonsson, 2013; Smith & Lipnevich, 2018) have all emphasized the dubious methodological quality and limited scope of studies that were used to examine student perceptions of feedback. Jonsson (2013) showed that the majority of his 103 reviewed studies into learners' use of feedback utilized surveys and interviews, and only two studies employed in vivo think-aloud protocols to capture student perceptions and interactions with feedback. Smith and Lipnevich (2018) discussed the dilemma that scholars face in studying feedback: whether researchers should observe without interference with the process of feedback provision and utilization, or whether scholars should implement more interventions. The conclusion that the authors provided is that the field could use both. There should be studies taking place in real classrooms involving some experimental manipulation that would allow for less ambiguity in the interpretation of results, but at the same time, we should try and not make the setting of the study as controlled as to stifle the generalizability of the findings. This conclusion may well apply to research into student perceptions.

1.5 The current review

Given concerns over the methodological quality and scope of research on student perceptions of feedback (e.g. Jonsson, 2013; Smith & Lipnevich, 2018), there is a need to review which methods have been employed in such research and what insights can be gained from such investigations. An overview of methods is needed to provide insights into the state of the field, to encourage researchers to conduct methodologically sound studies that add meaningful insights to the existing body of literature on this topic.

This scoping review (Arksey & O'Malley, 2005; Gough, Oliver, & Thomas, 2012; Peters et al., 2015) of the literature on students' perceptions of feedback aimed to (1) provide an overview of studies that have investigated student feedback perceptions, document the methods these studies employed, and identify strengths and limitations of these methods and (2) summarize main research findings to identify potential gaps in our understanding of student perceptions of feedback. These two main goals allowed us to report on the types of evidence currently available and what can be learned from research published thus far, to generate a set of recommendations for the future of the field. Our research questions were as follows:

1. How have students' perceptions of assessment feedback been investigated in studies published up to 2018, and what are the strengths and limitations of these methods?
2. What insights can be gained from the literature on student perceptions of feedback published up to 2018?

2 Method

2.1 Procedures

Given the nature of the aims and research questions of this review, a scoping review approach was considered most appropriate (Arksey & O'Malley, 2005; Gough et al., 2012; Peters et al., 2015). Sometimes called 'mapping reviews', scoping reviews are a relatively new approach to reviewing literature (Peters et al., 2015; Pham et al., 2014). According to Pham et al. (2014), there are three main differences between scoping reviews and systematic reviews: (1) scoping reviews aim to provide an overview of diverse body of literature on a particular topic, whereas systematic reviews aim to synthesize the best available evidence in relation to a highly specific question; (2) scoping reviews include a broad range of study designs, whereas systematic reviews often focus on research designs that allow for claims regarding effectiveness; and (3) scoping reviews aim to generate broad descriptive overviews, without synthesis of evidence from each individual study, whereas systematic reviews aim to critically synthesize evidence from a smaller number of studies. However, there appears to be much variation in the extent and nature of published scoping reviews (Arksey & O'Malley, 2005; Pham et al., 2014). For example, Arksey and O'Malley (2005) suggested that one common reason to undertake a scoping review is to identify gaps in the literature, which does require researchers to synthesize evidence about what is known to date, which is in contrast to Pham et al.'s (2014) third area of difference. At the same time, unlike systematic reviews, scoping reviews do not assess the quality of evidence in the reviewed studies with the aim to aggregate research findings based on the best available research evidence (Arksey & O'Malley, 2005). Thus, the nature of scoping reviews may vary considerably depending on the research aims and research questions the researchers seek to address (Arksey & O'Malley, 2005; Peters et al., 2015; Pham et al., 2014). We acknowledge that the nature of our research questions required analysis and synthesis beyond what is typical for scoping reviews (Arksey & O'Malley, 2005; Peters et al., 2015; Pham et al., 2014). For this reason, this review may be considered a 'critical scoping review'.

Although there appears to be substantial variation in methodological rigour of scoping reviews, such reviews generally appear to follow less stringent guidelines compared with systematic reviews (Pham et al., 2014). To ensure methodological rigour and replicability of the review, the procedures for this critical scoping review not only draw on the Joanne Briggs Institute guidelines (Peters et al., 2015) but also include elements of classic systematic reviews as described by Petticrew and Roberts (2006) and Gough et al. (2012). The critical scoping review included the following stages: formulation of the review purpose; trialling of search terms and databases; conducting the literature search; formulating inclusion criteria and selecting relevant literature using these criteria, extraction of data, charting and analysing data, and reporting the results. A data extraction form (see the Online Supplement) was designed to consistently record the same type of information for each of the selected publications.

2.2 Literature search

The literature search consisted of three rounds, described next. The three rounds of searching for literature and systematic selection of studies using the inclusion criteria (see Section 2.4) took place in iterative cycles.

2.2.1 Round 1

Following several trial searches, a literature database search using ERIC and PsycInfo was conducted mid-2017. The same search terms were used in both databases: the word 'feedback' had to occur in the title, and 'perception' or 'perspectives' (or variations on these terms) had to occur in the abstract.[1] No restrictions were placed on the publication date or publication type. In addition, further literature searches were undertaken through snowballing (White, 1994) using the selected literature. Snowballing involved scanning the reference lists of obtained sources for potentially relevant publications based on their titles, as recommended for use in scoping reviews by Peters et al. (2015).

2.2.2 Round 2

After completing the database and snowball searches, it was noted that several studies in primary and secondary education had not been found. A limitation of the snowball search is that it only allows for the identification of older sources. An additional 'successive snowballing' (Van der Kleij et al., 2019) search in Google scholar was therefore conducted to identify recent literature that may have been missed. The successive snowballing search involved searching for studies that had cited two widely cited articles on formative assessment and feedback in education: Hattie and Timperley (2007) and Black and Wiliam (1998). Many of the selected studies retrieved in round 1 reference either one of these or both these articles. For each of these articles, the full reference was entered into Google Scholar, and the list of citing studies was obtained by selecting 'cited by N', restricted to be published since 2011. The first 150 results (15 pages) were systematically scanned for relevant sources. A second snowball search was then conducted using the full-text versions of the additionally retrieved records.

2.2.3 Round 3

Finally, a second systematic literature database search was conducted in early 2019 using the same search terms and databases as those used in round 1, to ensure all relevant literature published between the mid-2017 search and the end of 2018 had been retrieved. This third round was necessary to provide an updated search (Pham et al., 2014), as reading and coding the considerable volume of publications retrieved in rounds 1 and 2 had taken 1.5 years. Given the updating purpose of the third search round, the search was restricted to literature published between 1 July 2017 and 31 December 2018. A third partial snowball search was conducted using the full-text versions of these studies. The snowball search for this round was only partial because the authors came to realize that it would not be feasible to include all relevant literature

[1] Search string: ti(feedback) AND ab(perception*) OR ti(feedback) AND perspect*

due to the rate at which new research was published. Further, saturation had been achieved, as other potentially relevant studies did not add new insights informative to the two research questions (Levinsson & Prøitz, 2017).

2.3 Inclusion criteria

Retrieved literature was exported to Thomson Reuters Endnote X7 (2013) for systematic screening and selection using the inclusion criteria. The inclusion criteria for scoping reviews are generally not highly restrictive, as the aim is to generate an overview of existing evidence, rather than answering very specific questions based on quality evidence (Arksey & O'Malley, 2005; JBI, 2015). Following screening based on study title, abstract, and keywords, studies were subjected to the following three inclusion criteria sequentially, in order to allow efficient identification of relevant literature:

- Criterion 1: The study focuses on students' perceptions of feedback.
- Criterion 2: The study reports empirical research.
- Criterion 3: The study was published in English.

The first criterion indicates that the research was conducted in an education setting and involved students. No restrictions were placed on the level of education or discipline area. Studies that focused on comparing teacher and student perceptions of feedback were also included. In this selection step, many irrelevant studies that related to, for example, students' perceptions of something other than feedback, but for which feedback was gathered from students, were excluded. Also, studies that merely focused on students' perceptions of feedback management systems, such as online submission systems were removed from the selection. Criterion 2 implied that only empirical research was considered, and purely descriptive studies or literature reviews were not included. However, no further restrictions were placed on the types of study designs. Criterion 3 was used to ensure that only highly accessible literature was included.

2.4 Data extraction and data analysis

A data extraction form was designed to suit the purpose of this critical scoping review (see the Online Supplement). The form was trialled with different coders and refined several times before all selected studies were coded using the final data extraction form.

To minimize bias in data extraction, all studies were coded using the data extraction form by the first author, and over 30% was blindly double coded by the second author or one of three trained research assistants. Although blind double coding is generally not required in scoping reviews (Arksey & O'Malley, 2005; Pham et al., 2014), this process was followed to ensure rigour in coding as typically required in systematic reviews (Gough et al., 2012). Inter-coder agreement was checked intermediately, and only minor differences were identified in wording used or amounts of information recorded. As many of the fields in the coding form did not require coders to make subjective judgements, and no major differences were identified in coding forms completed by different coders, double blind coding of 30% of the studies was deemed sufficient. For double-coded studies, both forms were used for analysis, and any

differences in coding were discussed between the authors and resolved by checking the full-text publications.

In order to answer research question 1, first, a descriptive overview of all included studies was generated, which contained the country in which the study was conducted, education level, subject or area of study, sample size(s), and type of publication (Table S1, online only). Based on the first author's reading of all publications, a draft framework for classifying publications according to their data collection methods was generated. To further ensure inter-coder agreement in analysing the results, population of this framework was completed by the second author for *all* studies, using double-coded data extraction forms when available, and the full-text versions of publications. Further, a research assistant coded 30% of studies by identifying methods and level of analyses. Inter-rater agreement was 98%. In addition to providing an overview of the reviewed studies' data collection methods, analysis involved the critical evaluation of the strength and limitations of each of these methods following standard methodological guidelines (see Gall, Gall, & Borg, 2007).

Main findings reported in the selected studies were examined to answer research question 2. This analysis was configurative in nature (Gough et al., 2012; Levinsson & Prøitz, 2017) and aimed to show the extent of evidence, as well as (the lack of) cumulative evidence within (sub)themes. However, no attempts were made to evaluate the strength of evidence, as is typical for aggregative reviews (Gough et al., 2012).

Analysis drew on the reported research findings and generated main themes and conclusions. Draft (sub)themes were inductively developed by firstly scanning the research aims and key findings for all studies, as displayed in Table S2 (online only). Next, the findings as listed in Table S2 were thematically coded using Nvivo 11. We did not intend to provide an unbiased exhaustive overview of all study findings as is typical for aggregative reviews (Gough et al., 2012). Rather, we summarized the main themes identified across the reviewed compendium of literature. Due to the large number of reviewed studies and a range of covered topics, we chose the Lipnevich et al. (2016) model as an organizing framework for narratively presenting the findings (Arksey & O'Malley, 2005; Gough et al., 2012). This student–feedback interaction model (Lipnevich et al., 2016) outlines characteristics of feedback, student, affective and cognitive responses, as well as actions that students take, all occurring in a specific context.

3 Results

3.1 Literature search and selection

In total, 164 studies identified in the literature search met the inclusion criteria. The detailed results of the literature search and selection process are provided in Table 1, reported using the elements within the PRISMA diagram (Moher, Liberati, Tetzlaff, & Altman, 2007), as recommended by Peters et al. (2015). The screening stage as shown in Table 1 reflects scanning publications for relevance based on their title, abstract, and keywords. Systematic screening and selection using the inclusion criteria is shown in Table 1 under eligibility. It is interesting to note that although the initial systematic data base search (round 1) resulted in almost half of the final, included studies, the

snowballing technique added another 29%. Also, the second systematic search (round 3), looking at just the most recent work, shows that this is a very active area of research at the present time.

3.2 Descriptive results

An overview of the included studies and their characteristics is provided in Table S1 (online only). The reviewed studies had been conducted in 52 different countries, with studies from the USA ($n = 33$, 29.5%) and Australia ($n = 15$, 13.4%) most well-represented. Over two thirds of the studies ($n = 128$, 75.3%) had been conducted in a tertiary education context, with small numbers of studies originating from secondary ($n = 25$, 14.7%), primary ($n = 15$, 8.8%), or non-education ($n = 2$, 1.2%) contexts. The studies had been conducted across a range of subject areas and/or disciplines. Among the different disciplines, foreign languages were most well represented ($n = 35$, 20%). Most studies ($n = 149$, 92%) had been published as a journal article. Of all journal article publications, the most widely represented journal was *Assessment and Evaluation in Higher Education*, with 26 articles (17.5% of all journal articles). Most studies

Table 1 Overview of literature search and selection results

Stage	Round 1	Round 2	Round 3	Snowballing
Identification	Systematic database search (mid-2017) •554 records identified through ERIC •1311 records identified through PsycInfo 1617 records after duplicates removed	12 records published since 2011 identified through citing 2 widely cited sources through Google Scholar	Systematic database search (mid-2017 and 2018) •57 records identified through ERIC •38 records identified through PsycInfo 91 records after duplicates removed	51 records identified through snowballing
Screening	Title, abstract, and keyword scan: 1517 records excluded	Title, abstract, and keyword scan: 6 records excluded	Title, abstract, and keyword scan: 61 records excluded	
Eligibility	100 full-text versions assessed for eligibility •3 publications excluded: no full-text available •16 publications excluded: did not meet inclusion criteria	7 full-text versions assessed for eligibility	30 full-text versions assessed for eligibility •1 publication excluded: did not meet inclusion criteria	51 full-text versions assessed for eligibility •4 publications excluded: did not meet inclusion criteria
Included	81	6	29	47
Total included	164			

only involved students, whereas 50 studies (30.5%) also involved teachers. The vast majority of studies focused on teacher feedback ($n = 152$), and a small number of studies (also) focused on peer feedback ($n = 15$). The number of student participants ranged from two to 2711. The number of teacher participants ranged from one to 460. Only a few studies had been published prior to the year 2000. Figure 1 shows the rapid increase in the number of publications on student feedback perceptions in recent years.

3.3 How students' feedback perceptions have been investigated

One of the main goals in writing this review was to examine methodological approaches that researchers had used to capture student perceptions of feedback and to discuss the strengths and limitations of these methods. Table 2 presents a summary of methodological approaches used to study student perceptions of feedback, along with the type of feedback towards which perceptions were gauged (i.e. level of generality, source). It was not our goal to critique each individual study and point at specific methodological limitations or highlight their strengths. We will, however, discuss each method, summarize common issues, and describe particularly prominent examples.

Surveys Table 2 shows that out of 164 studies 91 used surveys to capture student perceptions of feedback. In most of these studies, researcher-constructed survey instruments were used, although some used existing measures, not all of which have gone through a rigorous process of survey validation. Methodological advantages associated with survey research include opportunities to work with large samples, the power to generalize findings, especially when a representative sample is drawn from a well-defined population, and a relative ease of administration (Brown & Harris, 2018; Creswell & Guetterman, 2019). However, the quality of data and conclusions depends greatly on the sample as well as whether the instrument is rigorously constructed and exhibits good psychometric characteristics.

Not all studies included in our review that used a survey approach to gather student perceptions of feedback satisfied even moderately stringent criteria for survey

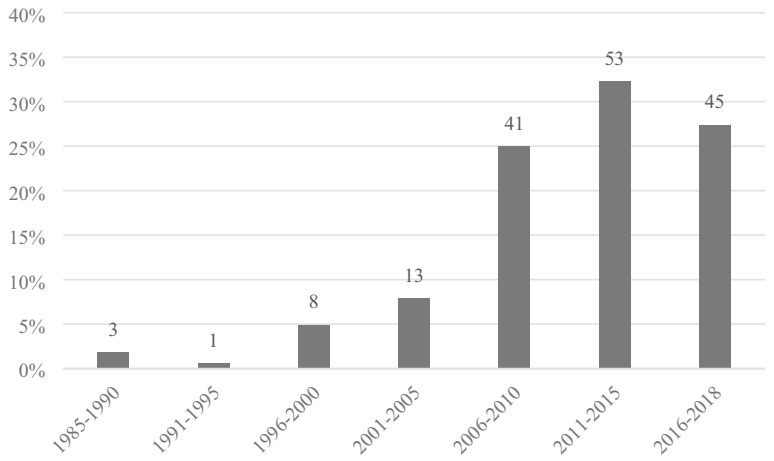

Fig. 1 Publication years of reviewed studies

Table 2 Methods employed in reviewed studies

Method	Number of studies
Surveys	91
Interviews	64
Focus groups	31
Open-ended questions/reflection prompts	25
Experimental manipulation	12
Observations	5
Think-aloud protocols	4
Level of generality	
Generic	108
Specific prompts	44
Immediate perceptions	12

Some studies used multiple approaches; hence, the totals exceed 164

construction and analysis (Raosoft, 2017). The main issues with studies that employed this method were (1) using insufficient samples for the analyses conducted, (2) devising questionnaires but not going through the process of survey validation, (3) not reporting psychometric characteristics of the instrument, (4) not including meaningful outcome variables when gauging student perceptions, and (5) not presenting proper description of the method or the analyses.

For example, Evans and Waring (2011) developed a questionnaire and piloted it in a sample of 20 students. The authors administered the survey to the same 20 students 2 weeks later to establish reliability evidence. The survey did not go through a systematic validation procedure, and the sample size of 20 students was far too small to make any meaningful conclusions regarding psychometric qualities. As a result, any conclusions derived from this study are questionable at the very best. Bayerlein's (2014) study provides another example. The author based their analyses on two items that came from a standard, anonymous, and university-wide course evaluation questionnaire. 'The feedback I received was provided in time to help me improve' was intended to index student perceptions of feedback timeliness, whereas 'I received constructive feedback on my work' was used to gauge student perceptions of the constructive nature of feedback. The analyses are not well described, but correlations are reported. Single-item assessments have been used for gauging affective states (Goetz, Pekrun, Hall, & Haag, 2006), but validity of single-item assessments in feedback research has yet to be established. Overall, collecting data on two items of dubious quality is not an advisable methodological choice.

There were also strong examples of survey application among the 164 reviewed studies. For example, King, Schrodt, and Weisel (2009), over the course of two studies, presented their attempt to develop and validate a survey that aimed at capturing student perceptions of teacher feedback. The authors carefully formulated and selected their items, used appropriate analytical techniques to establish and replicate the structure of the survey, and used meaningful outcomes—all of this in samples of over 200 students. In sum, there can be place for survey research in the sub-field of student perceptions of

feedback; however, scholars should approach it with greater care and follow 'industry' standards in how to design, validate, adapt, and use questionnaires to answer a variety of important research questions.

Interviews Interviews represent a direct way of collecting information about student perceptions of feedback and allow researchers to collect in-depth, often idiosyncratic, and insightful responses that surveys may not gather. Out of 164 studies included in our review, 64 reported interviews as their data collection method. Over 40 of these studies included serious methodological violations. The main problems that we observed in these studies were (1) small samples, that included as few as two students; (2) poorly described analyses; (3) poor or lacking description of the criteria for participant selection; (4) overstatement of results; and (5) not acknowledging the potential influence of the interviewer, especially in studies where the course instructor interviewed their students.

For example, Mustafa (2012), in an attempt to gather student perceptions of teacher feedback in Saudi Arabia, conducted an interview study with five students. The researcher first employed 'conversational interviews' that were followed by a second round of more structured interviews. Interviews, coding, analyses, and report preparation were conducted by the same person, which makes the results of the analysis somewhat questionable. The criteria for participant selection were not discussed, and coding and subsequent analyses were not explicated.

Interviews that were based on specific feedback that students received were more informative. For example, Hargreaves (2013) used videos of classroom observations to trigger student responses to feedback. The researcher video recorded nine students over time and at the end of each filming day asked the nine participants to comment on the specific feedback they had received. Stimulated recall studies provide stimuli for conversations, thus reducing the memory load and freeing up students' cognitive resources. Students can then invest these resources into considering their perceptions of each form of feedback. Among interview studies, only seven included this approach.

Focus groups Thirty-one studies included in the current review reported focus groups. One of the benefits of focus groups over one-on-one interviews includes larger sample size, faster data collection, and reduced influence of the moderator. However, whereas some claim focus group may elicit richer responses due to different perspectives and prior knowledge of participants (Brown & Harris, 2018), others have argued the opposite claiming higher cognitive load and potential production blocking and other cognitive biases (e.g. Nijstad & Stroebe, 2006). Other methodological shortcomings are similar to those observed for one-on-one interviews. Eva et al. (2012) conducted an impressive research in which they ran 17 focus groups with 134 participants. The researchers approached data collection and analyses with commendable rigour, working in teams of five with each transcript and deriving categories iteratively.

An example of a study that combined surveys and focus groups was that of Crimmins et al. (2016). The authors presented results of a survey administered to a large sample of students ($n = 414$ and $n = 353$, for two waves of data collection). The survey component included a total of four questions, with three of them being dichotomous yes/no items. The fourth item prompted students to report their satisfaction with feedback on a scale from (1) excellent to (5) poor. A teacher survey

component was similar in nature and included five items, three of which were dichotomous yes/no items, and the remaining three contained response categories that were not well defined (e.g. teachers were prompted to report on their emotions, and the choice of affective states was not supported by any theory). The analyses of the survey data included reported frequencies. The results of the survey component of the study are not very informative in that the quality of items is rather low and the dichotomous items cannot capture enough variance in participants' responses. Further, basic frequency analysis is not particularly helpful in moving the collective thinking forward. This study also included focus groups with eight students. The researchers fail to explain how exactly participant responses were coded and whether any form of triangulation was employed to ensure consistency of coding. This suggests that combining methods may not be ideal if each of the multiple methods employed is not utilized in accordance with strict methodological guidelines.

Open-ended written questions or reflection prompts In 25 out of 164 studies the researchers gathered data by asking participants to provide written responses to open-ended questions or prompts. Common methodological shortcomings across these studies included (1) small sample sizes, (2) limited number of prompts or questions, (3) weak justification for coding strategies, and (4) limited detail on how the analyses were conducted. Chanock (2000), for example, used open-ended items to assess participants' understanding of specific feedback phrases. The feedback was not provided on students' own work; rather, the researcher wanted to know how students perceived advice to include more analysis and present less description in their essays. The researcher did not report any detailed analyses and the description of the procedure, criteria for inclusion and exclusion, and other key details were highly ambiguous. Another example of non-optimal use of this method is a study by Douglas, Salter, Iglesias, Dowlman, and Eri (2016). In a sample of 587 students, Douglas et al. (2016) gathered participants' responses to feedback using a series of written open-ended questions. The study has an advantage of a large sample size, but there was a disconnect between the data collection method used and the data analysis employed. The researchers converted all qualitative data into quantitative (i.e. percentages) and thus reduced the richness of the data at hand at the same introducing potential subjectivity to coding and interpreting of the qualitative information.

In general, having students describe their thoughts about feedback in writing may be a very promising approach. It might reduce cognitive load present during focus group discussions (*brainwriting*, see Paulus & Yang, 2000) and shorten the time needed to conduct individual interviews, at the same time producing richer narrative. The results, however, depend greatly on the quality of prompts, coding schemes, and analyses.

Experimental manipulations Only 12 studies used experimental manipulations. Experimental studies allow for assessing effects of feedback on any set of possible student outcomes (cognitive, affective, or behavioural). For experimental findings to be robust, certain requirements have to be satisfied. These include (1) random assignment, (2) control groups, (3) inclusion of multiple dependent measures, and (4) sufficient sample size to carry out comparisons. An example of an experimental investigation includes a study by Clark-Gordon, Bowman, Watts, Banks, and Knight (2018) who investigated the impact of (1) the presence or absence of an instructor picture on perceptions of

feedback and the instructor and (2) feedback including emojis on feedback perceptions. The authors found that face-threat mitigation positively affected student perceptions of the instructor and their feedback and enhanced students' sense of social presence. Furthermore, presence of the instructor's picture and emojis did not enhance students' sense of social presence. Similarly, Pitt and Winstone (2018) examined students' perceptions of fairness of anonymous vs non-anonymous marking, and whether these are related to student gender and ethnicity, and students' perceptions of the relational aspects of non-anonymous marking. The analyses revealed no differences in student achievement following anonymous or non-anonymous marking and no overall differences in perceived fairness of anonymous or non-anonymous marking (although women perceived anonymous marking to be fairer). Students viewed feedback from non-anonymous markers as more helpful, because it allowed them to request verification, seek further feedback, and encouraged them to improve. Findings of this nature would be difficult to obtain in a non-experimental study. After all, experiments are among the few approaches that allow for causal conclusions. Such studies provide evidence-based insights into student perceptions of feedback across distinct conditions, which has substantial potential to inform practice. For example, it is much more informative to know that there is no difference in students' perceptions of fairness for anonymous vs non-anonymous marking than knowing that students thought either of the two were fair.

Some studies that included experimental manipulations had design flaws that made their conclusions problematic. For example, McGrath, Taylor, and Pychyl (2011) examined student perceptions of feedback depending on its type. The procedures are not well described, the sample size is far too small, and not enough details are presented for the reader to be able to evaluate the robustness of the conclusions. In sum, experiments are only useful when basic methodological guidelines are followed in conducting them.

Observations and think-aloud protocols Out of 164 studies, four and five studies employed think-aloud protocols and observations, respectively. Think-aloud protocols require individuals to complete a task while explaining their thoughts and reactions to it (Brown & Harris, 2018). This information may prove very useful, as student perceptions of specific feedback messages can be revealed. Unfortunately, very few studies resorted to this methodological approach. For example, Mahfoodh and Pandian (2011) investigated students' affective reactions to written teacher feedback along with contextual factors influencing student perceptions of feedback. Students were asked to read commented drafts and describe their immediate take on the feedback that they receive. The challenge with think-aloud is in student ability to activate their meta-cognitive skills and share their thoughts and emotions while working on a task. This requires training and effective facilitation.

The five studies that used observations included this method as a supplement to the primary method. The strength and weakness of observations are in its relatively unstructured nature. The participants are observed in a natural setting, without any interference from researchers. Researchers have some control over whom they are going to observe and which behaviours they are going to record. Needless to say, the result of any observation depends on the visibility and sufficient presence of the behaviour of interest. There is also the potential of researcher impact on results.

Observations serve as a good starting point for other methods, such as interviews and work sample analysis. For example, Nguyen and Filipi (2018) and Murtagh (2014) used observations along with focus groups, interviews, and reviews of student work.

Level of generality In coding our studies, we also focused on the generality of feedback. Lipnevich et al. (2016) proposed that students' feedback perceptions are made up of a two-level structure: (a) state, involving situation-specific, perceptual, and automatic responses and (b) trait, involving 'typical', reflective, and controlled responses. It appears that the majority of the research on student feedback perceptions has been focused on measuring 'trait' level perceptions, elicited by surveys or group/individual interviews. One hundred eight studies in our sample examined student perceptions of feedback in general. In other words, students were asked to imagine a type of feedback or engage in recall of averaged feedback episodes from their school/academic career. Conversely, 44 studies presented specific feedback and asked students to report their reactions to this particular type of feedback. A small subset of studies—12—had students respond to feedback specific to their work immediately following feedback. Only the latter subset can capture so called 'state' perceptions of feedback, and there is clearly a skew favouring generic feedback.

3.4 Insights gained from student feedback perception literature

This section presents the main insights gained from examining the 164 studies (see Table S1, online only) on student feedback perceptions. We will present our findings as aligned with the Lipnevich et al. (2016) model, discussing student perceptions of feedback depending on its type, summarizing student characteristics that affect perceptions, and presenting affective and cognitive responses that are linked to or translate into student perceptions, across contexts. For a more detailed summary of findings, see Table S2 (online only). The numbers reported in the results section correspond with those in Tables S1 and S2 (online only).

3.4.1 Characteristics of feedback

Usefulness and effectiveness of feedback were two of the most studied characteristics of feedback. Several studies that focused on perceived usefulness of feedback showed that students viewed feedback as useful (16, 43, 51, 87, 95, 158, 160) whereas others revealed that students perceived feedback to be potentially useful (40, 154) or not useful at all (22, 38, 88, 137). Several studies investigated student perceptions of feedback quality (14, 16, 100, 106, 49, 150, 151, 152). There was not a common theoretical framework underlying these inquiries, so student perceptions of features that made feedback useful or high quality varied significantly across studies. For example, three studies identified that students did not see utility in feedback that was too generic and did not directly apply to their individual work (128, 154, 160). Interestingly, only one study provided evidence that perceived usefulness of feedback positively related to student achievement and interest (63).

Another well-studied characteristic of feedback was its perceived effectiveness or quality (2, 17, 54, 56, 63, 90, 114, 123, 124, 136, 140, 145, 147, 148, 155, 156, 160,

161). There was limited consistency in feedback features identified as effective across studies, except for corrective feedback (16, 35, 76), the need for feedback to occur in interactions between teachers and students (45, 117, 120), and the need for feedback timeliness (56, 57, 68).

Mode of feedback provision was the focus of several investigations. Studies reported students' positive perceptions of audio feedback (60, 92, 125), feedback provided through multiple channels (written and verbal, provided individually and in a group setting), answer-until-correct feedback systems (136), or feedback presented through online submission and marking systems (152). Interestingly, none of these studies provided any evidence of the actual effectiveness of these feedback characteristics.

Another characteristic of feedback that received attention from scholars is the source of feedback. For peer feedback, studies found that students perceived providing and/or receiving peer feedback as useful (11, 73, 86, 129, 130, 134, 135), whereas three studies reported that students' opinions regarding the value of peer feedback varied (47, 128, 146). Furthermore, three studies suggested that the peer's competence level affected peer feedback perceptions (15, 37, 140), and two studies identified a safe classroom climate as a prerequisite to useful peer feedback (135, 141). Once again, none of the reported studies linked student perceptions of peer feedback to meaningful outcomes.

Studies also examined whether or not there was a clear alignment between student and teacher perceptions of feedback. Fewer studies reported perceived alignment between teacher and student perceptions of feedback (good alignment: 10, 13, 16, 79, 87, 98, 126, 160; some alignment: 8, 22, 77, 85, 101, 143, 149) compared with studies that reported differences in perceptions of (specific aspects of) feedback between students and teachers (1, 8, 13, 14, 22, 25, 34, 36, 39, 59, 65, 68, 77, 85, 93, 94, 98, 101, 102, 120, 127, 144, 159, 161, 162). For example, several studies showed that teachers thought their feedback was more useful than students did (22, 65, 101). The reasons behind differing teacher and student perceptions varied across studies. These reasons included differences in perceptions related to feedback as an ongoing process versus something that occurs following assessment (1), motives for feedback provision (8), the extent of error correction (13), ideas about the purpose of assessment (93), dissatisfaction with particular aspects of feedback practices (94), and unrealistic student expectations (102). Findings regarding perceived purposes of feedback were also inconsistent across studies (16, 93, 110, 111, 114, 127, 145).

Several studies highlighted the importance of social-relational aspects in feedback exchanges, including interpersonal communication (22, 45, 75, 85, 117, 120, 149, 159). For example, several studies highlighted the strengths of dialogic forms of feedback (75, 117, 120, 149), noted the importance of interpersonal relationships (38, 54, 81, 135, 139, 141) in facilitating productive feedback practices, and examined how power relationships between the different parties in feedback interactions impact feedback perceptions (18, 22, 23, 61, 88, 149).

Other aspects of feedback characteristics that affected reported student perceptions included feedback timing (56, 57, 58, 137, 148, 150, 152) feedback source (12, 33, 116, 128, 130), amount (16, 137, 150, 160), specificity (16, 62, 128), or valence (i.e. positive vs negative feedback; 29). Several studies (36, 82, 91, 139, 164) described student feedback perceptions on a generic level, without specifying any of the identified feedback characteristics.

3.4.2 Student characteristics

Studies examined student characteristics that influenced or moderated student perceptions of feedback. Student perceptions of feedback were moderated by students' year of study (3, 4, 19, 118, 132, 137) and age (17, 27), although the evidence was inconclusive regarding how these demographic variables influenced perceptions. Some studies suggested that student perceptions of feedback became more negative as they progressed through their studies (2, 4, 137), whereas one study suggested that final year students' feedback perceptions were more nuanced and aligned to the literature on effective feedback (118). Other demographic variables included whether students were international or domestic (132).

Further, eight studies identified gender as an important factor in determining student feedback perceptions (2, 26, 65, 107, 108, 116, 132, 156), but in relation to different aspects such as fairness (116), quantity, feedback type (26, 107, and 108), or satisfaction (132), with no consistent findings across studies. Two studies found that students' feedback perceptions partially mediated the link between student self-efficacy and self-regulation (44, 164), with another study suggesting that confidence determined how feedback was perceived (46). Other aspects identified by several studies as (not) influencing student feedback perceptions included affective variables (22, 39, 46), individual student differences (88, 156, 162), student academic attainment (88, 98, and 2 found there was no relationship), students' goal orientation (30, 123, and 55 found there was no relationship), self-calibration (55), or epistemological beliefs (109).

3.4.3 Student responses and context

Study findings regarding student behavioural, cognitive, and affective responses to feedback were variable. Whereas some studies reported that students used feedback (105, 111, 128, 147) directly or in follow-up tasks (110, 131), others reported limited student use of feedback (66, 93, 127), perhaps because students lacked the skills for using the feedback effectively (48, 88, 111). Several studies pointed out that how students use feedback will depend strongly on individual student differences (75, 116, 155). Feedback mode did not appear to influence the quality of students' feedback use (32), and time spent reading peer feedback did not relate to extent of feedback use (15).

Various studies highlighted the importance of considering emotions when examining student feedback perceptions. Emotions were shown to positively or negatively relate to how feedback is perceived (39, 46, 55, 62, 64, 90, 95, 96, 97, 99, 105, 115, 127, 131, 140, 143, 144, 147, 160), with studies reporting negative affective reactions in situations where students did not understand feedback (7, 95).

Further, several studies revealed that, overall, students chose to pay attention to feedback (43, 51, 95, 127, 131). The extent of attention paid to feedback depended on the type of feedback information (43) or the stage in the assessment process (51). One study reported that some students avoided receiving feedback (97). Attention paid to incorrectly answered items was found to be associated with enhanced achievement in one study (53). Various studies highlighted a lack of understanding of feedback as a barrier to engagement with feedback (7, 25, 34, 46, 51, 95, 105, 154), whereas two studies identified that students felt that they could accurately interpret the meaning of teacher feedback (29, 87).

Studies identified a range of student variables influencing student active engagement with feedback, including individual student differences (55, 75, 111, 115, 145), students' capacity to self-regulate their learning (52, 55, 102, 142), willingness to learn from feedback (15, 37, 48, 73), gender (108, 133, 156), levels of student autonomy (142), and ability, persistence, and levels of confidence (52). Two studies (3, 4, based on the same dataset) identified that student engagement with feedback declined as students progressed through their studies. One study reported that one-on-one conversations encouraged student engagement, as it enables students to request immediate clarification of feedback (81). Only two studies identified that students actively sought feedback (72, 124).

In general, across studies, only a small number of researchers examined links between students' feedback perceptions and achievement outcomes. Quite surprisingly, only one study (53) identified that how students engaged with feedback predicted their achievement levels. Although some reported positive relations between student perceptions of feedback and achievement (31) or perceptions of perceived usefulness and student achievement (63, 107, 123), a study suggested that student perceptions of the accuracy of peer feedback did not relate to their willingness to use such feedback or actual improvements made to their work (73). Similarly, results of another investigation revealed that students' perceptions of feedback usefulness did not relate to their performance (148). Further, several studies linked positive student feedback perceptions to self-reported improvements in student learning (106, 125, 129, 136, 158), or negative perceptions of feedback to a self-reported lack of learning (69, 105, 145).

In terms of contextual variables, nine studies found that the opportunity for dialogues positively influenced students' feedback perceptions (4, 14, 18, 24, 31, 65, 120, 150, 163). Five studies suggested that the assessment context impacted on student perceptions of feedback (9, 80, 88, 133, 144). Other studies found that student feedback perceptions were influenced by teacher factors such as their expectancy level of students or pedagogic approach (26, 28, 88, 98), guidance on how to use feedback (14, 137, 154), opportunity to use feedback (54, 124), credibility of the feedback provider (119, 133), or the extent of learning support (39).

All in all, the findings were quite scattered and although they did paint a picture of how students perceived feedback, a disproportionately small number of studies linked these findings to meaningful educational outcomes.

4 Discussion

This critical scoping review investigated (1) how students' perceptions of assessment feedback have been investigated in studies published between up to 2018 and what the strengths and limitations of these methods are and (2) what insights about student perceptions of feedback can be gained from this body of research. In the following sections, we discuss central findings of our study.

4.1 Methodological approaches in feedback perception research

The second main goal of this review was to examine methodological approaches that researchers utilize to examine student perceptions of feedback and to highlight the

strengths and limitations of these methods (see Gall, Gall, & Borg, 2007). In this section, we summarize our main findings and make recommendations for future research based on these findings.

4.1.1 Surveys

Over 50% of the reviewed studies used surveys to capture student perceptions of feedback. This finding is consistent with those of Brown and Harris (2018). In their review of methods used in feedback research, Brown and Harris (2018) mentioned the preponderance of studies employing surveys to capture participants' attitudes, understanding, and general feedback experiences. Unfortunately, the overall rigour of survey methodology in the selected studies was poor. The main limitations of survey research included small samples, poor item quality, lacking evidence of adequate procedures of instrument validation, and poor description of data analysis. Future studies employing survey research methodology should resort to this method only if they can avoid the aforementioned limitations.

4.1.2 Focus groups

Focus groups were frequently used in our reviewed studies. There are varying opinions as to whether focus groups produce richer responses than interviews. Some researchers argue that focus groups elicit rich responses by allowing participants to talk to each other (Brown & Harris, 2018), whereas others claim that focus groups may introduce significant cognitive load and hence reduce the probability of obtaining rich responses (Nijstad & Stroebe, 2006). This may happen because group members often have to delay expressing their ideas until other participants have stopped talking. These speaking turn delays are the main cause of production blocking. As shown by Nijstad, Stroebe, and Lodewijkx (2003), production blocking is mainly due to cognitive interference: During speaking turn delays, group members have to monitor the discussion for opportunities to 'jump in'. This extra cognitive load may interfere with idea generation and may introduce anchors that prevent participants from retrieving meaningful and different ideas. Researchers should be mindful of these potential problems with this method when investigating student perceptions of feedback. Having clearly formulated research questions, a trained moderator, a script, or a prompt that can be used to guide the discussion would be highly advisable.

4.1.3 Experimental studies

Experimental studies with control groups and random assignment have been viewed as the gold standard of research. There were very few experimental studies among the 164 that we reviewed, and only few of them were exemplary. We acknowledge that educational settings are complex, so in some cases proper random assignment may not be possible. Thus, we would argue that conducting small scale in situ experimental studies may be of great importance to the field. This way, we will be able to attribute outcomes to experimental manipulation without sacrificing ecological validity and generalizability of findings (Smith & Lipnevich, 2018).

4.1.4 Level of generality

Most of the 164 studies in this review examined students' feedback perceptions at a generic level (Lipnevich et al., 2016), as opposed to 'state' perceptions. Research into achievement emotions has long differentiated between trait and state assessments (see Goetz et al., 2006; Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013). Drawing upon the emotions line of research, it is not too fanciful to speculate that trait assessments may strongly reflect overall cognitive structures about feedback rather than actual experiences, and they can be tainted by recall biases (e.g. Roseman, Antoniou, & Jose, 1996). The use of objects for stimulated recall, such as video, can help overcome such biases. Only a small number of studies had used such stimuli. In contrast, when using real-time assessments, such beliefs might play a less important role and focus on students' actual perceptions and not their generalized beliefs or memories about feedback. This assumption is based on dual-process approaches in higher cognition (Evans, 2010; Kahneman, 2003), with state assessments reflecting mainly processes of system 1 (fast, automatic, impulsive, perceptual) and trait assessments reflecting mainly processes of system 2 (slow, controlled, reflective, analytic). According to this perspective, state assessments represent the experiencing self (relatively independent of cognitions), whereas trait assessments reflect the remembering self (strongly impacted by cognitions; Kahneman, 2003). We are not trying to argue that one of these approaches is better than the other. However, researchers in the field of student perceptions of feedback might consider differentiating between the two levels of generality. Considering the preponderance of general-level feedback perceptions studies (93% of all reviewed studies), we would benefit from alternative, more momentary perspectives. In future research, methods not used in the studies in this review may be used to examine students' immediate feedback perceptions. For example, fMRI and eye tracking could be employed to examine students' cognitive and emotional reactions to feedback messages.

4.2 Key insights on student perceptions of feedback

Reviewing an entire sub-field of research is not an easy undertaking, and this work, ideally, should produce some meaningful findings that would inform both research and practice. Unfortunately, lacking theoretical frameworks, repetitiveness (not replicability) of studies, and methodological problems that we observed among the 164 included in this review have produced somewhat disappointing conclusions.

One key observation is that student perceptions of feedback is not a well-defined field, with isolated bodies of cumulative research occurring in different sub-fields, such as higher education and ESL education. However, even within these sub-fields, researchers have not drawn on common theoretical foundations. This might be a result of the ill-defined nature of the process of student engagement with feedback. Further, much of the research on student feedback perceptions is driven by practical concerns, such as student dissatisfaction or teacher frustration. For example, evidence from higher education student satisfaction surveys suggests students in the UK (Higher Education Funding Council for England, 2016) and Australia (Quality Indicators for Learning and Teaching, 2017) are generally unsatisfied with feedback and assessment practices. This

review contributes to providing a uniting overview of the feedback perceptions literature by combining insights from these sub-fields.

The general lack of common theoretical frameworks underpinning investigations has resulted in a wide range of examined variables. This diversity in variables makes it difficult to directly compare results of studies to generate useful insights. Even when studies focused on the same variables, results were often inconsistent across studies, which may be caused by differences in study populations, variable operationalizations, methods used, and rigour of the studies. Further, the findings within this body of research provide very limited insights into how student perceptions of feedback relate to engagement with feedback and subsequent meaningful outcomes. Hence, many fall short of answering the 'so what?' question. That is, if we know that students perceive a certain type of feedback as more useful, it does not necessarily mean that it does in fact translate into greater personal or academic outcomes. Moreover, Smits, Boon, Sluijsmans, and Van Gog (2008) argued that the kind of feedback that students perceive as useful and enjoyable may not always lead to best improvement. For example, students may prefer explicit corrections, but subtle hints requiring students to self-identify and correct mistakes may be more effective in the longer term. Without links to meaningful performance outcomes, the value of research into student feedback perceptions is somewhat questionable. Notably, those investigations that did link student perceptions to performance, revealed variable findings.

Based on this review's findings, Fig. 2 presents an overview of potential variables to consider when examining student perceptions of feedback. As indicated, the results showed variable strength and direction of evidence for each of these variables, and there may be other relevant variables to be considered. The identified variables are consistent with those in Lipnevich et al.'s (2016) model, although there are slight differences in wording. For example, in our review, the importance of accuracy of a feedback message was identified through the perceived competence of the feedback provider. Further, it must be acknowledged that this review did not seek to include literature specifically focused on feedback use (the action column in Fig. 2). However, the review conducted by Jonsson (2013) suggests that limited research has addressed this step in the feedback process, and there is a great need for empirical research into how students take action based on feedback. The framework presented in Fig. 2 can assist in future investigations of student feedback perceptions by highlighting relevant aspects to consider. Consistency in variables under investigation can assist in building a more robust evidence base in this field with scattered research findings.

4.3 Limitations and recommendations for future research

As is typical for scoping reviews (Pham et al., 2014) and review studies in general (Gough et al., 2012), a limitation of this review is that despite three rounds of literature searches, not all relevant literature may have been retrieved. However, as saturation had been achieved (Levinsson & Prøitz, 2017), it is unlikely that including further research published up to 2018 would have yielded different findings. Further, as previously noted, we acknowledge that the analysis and synthesis in our review went beyond what is typical for scoping reviews (Arksey & O'Malley, 2005; Peters et al., 2015; Pham et al., 2014). For this reason, we labelled the approach taken a 'critical scoping review'.

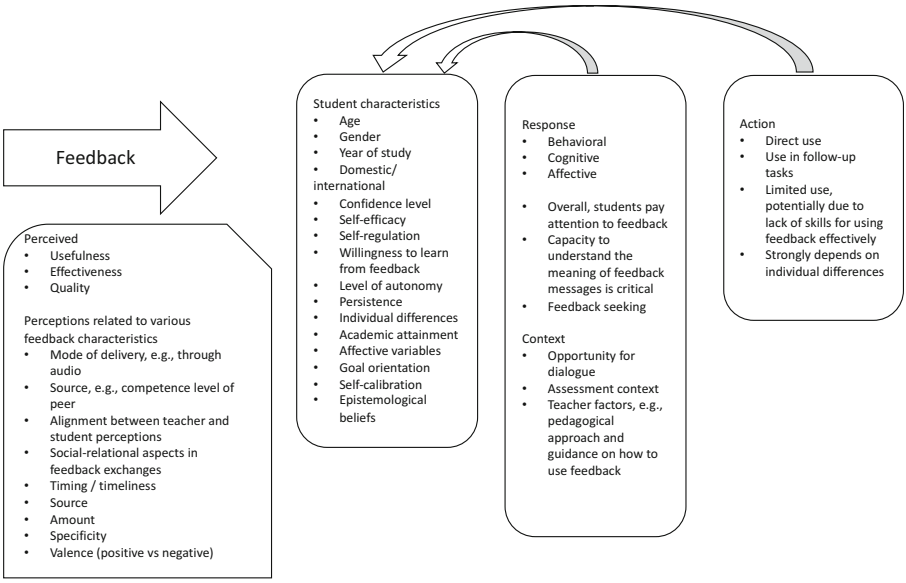

Fig. 2 Main findings as mapped onto Lipnevich et al. (2016) model

Due to the nature of this review, which did not involve assessment of the quality of evidence in synthesizing what insights can be gained from the literature on student perceptions of feedback (Arksey & O'Malley, 2005), the results for the second research question must be interpreted with some caution. Differences in methodological rigour of studies may explain some of the identified inconsistencies in findings. Most notable was the variability in findings in relation to student feedback perceptions and performance. Some studies showed that positive views on feedback linked to self-reported improvement of learning, whereas others did not. Similar results have been previously reported by Panadero (2016) and Alqassab and Panadero (2020), with researchers reporting lacking evidence of links between student perceptions of peer assessment and performance outcomes. It is quite possible that one of the potential explanations for a disproportionately low number of studies examining relations between perceptions of feedback and student performance can be explained by a file drawer problem (Bradley & Gupta, 1997). The latter rests on the assumption that statistically non-significant results are less likely to be submitted and published in journals. Meta-analyses can statistically mitigate this bias, which is virtually impossible to do in scoping or systematic reviews. We would like to echo calls coming from a variety of fields asking researchers to prepare and publish articles with non-significant findings and, whenever possible, to use the format of registered reports. The latter help to eliminate issues of selective reporting and publication bias and thus reward best scientific practices.

All in all, we recommend that in future investigations of feedback perception researchers should adhere to rigorous methodological guidelines to ensure the research will result in valid and reliable conclusions. Methods may also meaningfully complement each other, which some of the reviewed studies already showed (e.g. Harris, Brown, & Harnett, 2014). Furthermore, when preparing manuscripts, a detailed description of the sample, context, and procedures should be presented so that researchers, practitioners, and policy makers could judge whether the findings may or

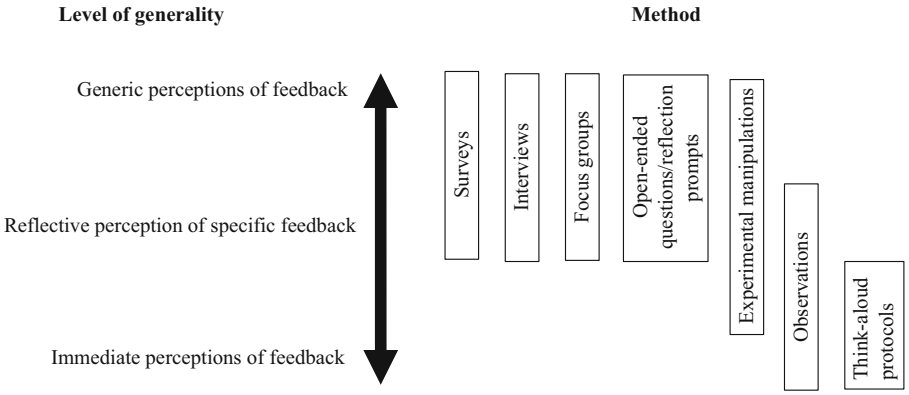

Fig. 3 Methods based on generality of feedback

may not help them to answer specific questions or solve problems at hand. As Wiliam (2019) notes, 'In educational research, "What works" is usually the wrong question because almost anything works somewhere, and nothing works everywhere. A better question is, "Under what circumstances does this work…"' (p. 11). Feedback perception research is not an exception: Describing specific contexts and domains, in which perceptions may link to improved outcomes, may be a fruitful area for future investigations. Comparing across cultures, developmental levels and academic domains may greatly inform the field and allow for more nuanced understanding of student views on feedback and how these might relate to scholastic achievement outcomes.

Based on the key findings of the critical scoping review and conceptualizations in the literature (Lipnevich et al., 2016; Van der Kleij & Adie, 2020), Fig. 3 presents a framework for future investigations into student perceptions of feedback. The aim of this framework is to assist researchers in considering the level of generality of feedback perceptions, along with appropriate methods to address the concept under investigation.

In sum, this study is the first critical scoping review of research into methods used in and key findings from research on student perceptions of feedback. This exercise revealed various issues that permeate this field of research, but at the same time offered promising avenues for future research. It is our hope that this critical scoping review will make researchers be aware of the many studies that already exist in this field and the various methodological approaches and prompt researchers to conduct studies that are methodologically sound and that investigate student perceptions of feedback within the nomological net of other related constructs, linking it to meaningful educational outcomes.

References

References marked with an asterisk indicate studies included in the critical scoping review. The full list of reviewed studies is provided in the online supplementary section.

Alqassab, M., & Panadero, E. (2020). Peer assessment. In S. M. Brookhart (Ed.), *Routledge encyclopedia of education*. New York: Routledge.

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology: Theory and Practice, 8*, 19–32. https://doi.org/10.1080/1364557032000119616.

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*, 213–238. https://doi.org/10.3102/00346543061002213.

Bayerlein, L. (2014). Students' feedback preferences: how do students react to timely and automatically generated assessment feedback? *Assessment & Evaluation in Higher Education, 39*, 916–931. https://doi.org/10.1080/02602938.2013.870531.

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

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*, 5–31. https://doi.org/10.1007/s11092-008-9068-5.

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment (UCLA-CSIEP), 1*(2), 1–12.

Bradley, M. T., & Gupta, R. D. (1997). Estimating the effect of the file drawer problem in meta-analysis. *Perceptual and Motor Skills, 85*(2), 719–722.

Brown, G. (2011). Self-regulation of assessment beliefs and attitudes: a review of the students' conceptions of assessment inventory. *Educational Psychology, 31*, 731–748. https://doi.org/10.1080/01443410.2011.599836.

Brown, G. T. L., & Harris, L. R. (2018). Methods in feedback research. In A. Lipnevich & J. Smith (Eds.), *The Cambridge handbook of instructional feedback (Cambridge handbooks in psychology)* (pp. 97–120). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316832134.021.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research, 65*, 245–281. https://doi.org/10.3102/00346543065003245.

*Carless, D. (2006) Differing perceptions in the feedback process. Studies in Higher Education 31, 219–233. doi:https://doi.org/10.1080/03075070600572132

Carless, D. (2016). Feedback as dialogue. *Encyclopedia of educational philosophy and theory*, 1–6.

*Chanock, K. (2000). Comments on essays: do students understand what tutors write? Teaching in Higher Education, 5, 95–105. doi:https://doi.org/10.1080/135625100114984

*Clark-Gordon, C. V., Bowman, N. D., Watts, E. R., Banks, J., & Knight, J. M. (2018). "As good as your word": face-threat mitigation and the use of instructor nonverbal cues on students' perceptions of digital feedback, Communication Education, 67, 206–225. doi:https://doi.org/10.1080/03634523.2018.1428759

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. New York: Addison Wesley.

*Crimmins, G., Nash, G., Oprescu, F., Liebergreen, M., Turley, J., Bond, R., & Dayton, J. (2016). A written, reflective and dialogic strategy for assessment feedback that can enhance student/teacher relationships. Assessment & Evaluation in Higher Education, 41, 141–153. doi:https://doi.org/10.1080/02602938.2014.986644

*Douglas, T., Salter, S., Iglesias, M., Dowlman, M., & Eri, R. (2016). The feedback process: perspectives of first and second year undergraduate students in the disciplines of education, health science and nursing. Journal of University Teaching and Learning Practice, 13(1). Retrieved from: http://ro.uow.edu.au/jutlp/vol13/iss1/3

*Eva, K. W., Armson, H., Holmboe, E., Lockyer, J., Loney, E., Mann, K., & Sargeant, J. (2012). Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. Advances in Health Sciences Education, 17, 15–26. doi:https://doi.org/10.1007/s10459-011-9290-7

Evans, J. S. B. (2010). Intuition and reasoning: a dual-process perspective. *Psychological Inquiry, 21*, 313–326. https://doi.org/10.1080/1047840X.2010.521057.

*Evans, C., & Waring, M. (2011). Student teacher assessment feedback preferences: the influence of cognitive styles and gender. Learning and Individual Differences, 26, 171–109. doi:https://doi.org/10.1016/j.lindif.2010.11.011

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: an introduction* (8th ed.). Boston: Pearson Education.

*Gamlem, S. M., & Smith, K. (2013). Student perceptions of classroom feedback. Assessment in Education: Principles, Policy & Practice, 20, 150–169. doi:https://doi.org/10.1080/0969594X.2012.749212

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science, 24*, 2079–2087. https://doi.org/10.1177/0956797613486989.

Goetz, T., Lipnevich, A., Krannich, M., & Gogol, K. (2018). Performance feedback and emotions. In A. Lipnevich & J. Smith (Eds.), *The Cambridge handbook of instructional feedback (Cambridge handbooks*

in psychology) (pp. 554–574). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316832134.027.

Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*, 289–308. https://doi.org/10.1348/000709905X42860.

Gough, D., Oliver, S., & Thomas, J. (2012). *An introduction to systematic reviews*. London: Sage.

Guskey, T. (2018). Feedback, correctives, and the use of pre-assessments. In A. Lipnevich & J. Smith (Eds.), *The Cambridge handbook of instructional feedback (Cambridge handbooks in psychology)* (pp. 432–450). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316832134.021.

Handley, K., Price, M., & Millar, J. (2011). Beyond "doing time": investigating the concept of student engagement with feedback. *Oxford Review of Education, 37*, 543–560. https://doi.org/10.1080/03054985.2011.604951.

Hargreaves, E. (2011). Teachers' feedback to pupils: "like so many bottles thrown out to sea"? In R. Berry & B. Adamson (Eds.), *Assessment reform in education: policy and practice* (pp. 121–133). Dordrecht: Springer.

*Hargreaves, E. (2013). Inquiring into children's experiences of teacher feedback: reconceptualising assessment for learning. Oxford Review of Education, 39, 229–246. doi:https://doi.org/10.1080/03054985.2013.787922

*Harris, L. R., Brown, G. T. L., & Harnett, J. A. (2014). Understanding classroom feedback practices: a study of New Zealand student experiences, perceptions, and emotional responses. Educational Assessment, Evaluation and Accountability, 26, 107–133. doi:https://doi.org/10.1007/s11092-013-9187-5

Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

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

*Havnes, A., Smith, K., Dysthe, O., & Ludvigsen, K. (2012). Formative assessment and feedback: making learning visible. Studies in Educational Evaluation, 38, 21–27. doi:https://doi.org/10.1016/j.stueduc.2012.04.001

Higher Education Funding Council for England. (2016). National student survey results 2016. Retrieved from http://www.hefce.ac.uk/lt/nss/results/2016/

Joanna Briggs Institute. (2015). *The Joanna Briggs institute reviewers' manual 2015: Methodology for JBI scoping reviews* (pp. 1–24). Joanne Briggs Institute.

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*, 63–76. https://doi.org/10.1177/1469787412467125.

Jonsson, A., & Panadero, E. (2018). Facilitating students' active engagement with feedback. In A. Lipnevich & J. Smith (Eds.), *The Cambridge handbook of instructional feedback (Cambridge handbooks in psychology)* (pp. 531–553). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316832134.026.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American Psychologist, 58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697.

*King, P. E., Schrodt, P., & Weisel, J. J. (2009). The Instructional Feedback Orientation Scale: conceptualizing and validating a new measure for assessing perceptions of instructional feedback. Communication Education, 58, 235–261. doi:https://doi.org/10.1080/03634520802515705

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254–284. https://doi.org/10.1037/0033-2909.119.2.254.

Levinsson, M., & Prøitz, T. S. (2017). The (non-)use of configurative reviews in education. *Education Inquiry, 8*, 209–231. https://doi.org/10.1080/20004508.2017.1297004.

Lipnevich, A. A., Berg, D. A. G., & Smith, J. K. (2016). Toward a model of student response to feedback. In G. T. L. Brown & L. R. Harris (Eds.), *The handbook of human and social conditions in assessment* (pp. 169–185). New York: Routledge.

*Lipnevich, A. A., & Smith, J. K. (2009). 'I really need feedback to learn:' students' perspectives on the effectiveness of the differential feedback messages. Educational Assessment, Evaluation and Accountability, 21, 347–367. doi:https://doi.org/10.1007/s11092-009-9082-2.

Lipnevich, A. A., & Smith, J. K. (Eds.). (2018). *The Cambridge handbook of instructional feedback*. Cambridge: Cambridge University Press.

*Mahfoodh, O. H. A., & Pandian, A. (2011). A qualitative case study of EFL students' affective reactions to and perceptions of their teachers' written feedback. English Language Teaching, 4(3), 14–25.

*McGrath, A. L., Taylor, A., & Pychyl, T. A. (2011). Writing helpful feedback: the influence of feedback type on students' perceptions and writing performance. The Canadian Journal for the Scholarship of Teaching and Learning, 2(2), 5. doi:https://doi.org/10.5206/cjsotl-rcacea.2011.2.5

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2007). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *PLoS Medicine,* 6. https://doi.org/10.1371/journal.pmed.1000097.

Mory, E. H. (1996). Feedback research. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 919–956). New York, NY: Macmillan Library Reference USA.

*Murtagh, L. (2014). The motivational paradox of feedback: teacher and student perceptions. The Curriculum Journal, 25, 516–541. doi:https://doi.org/10.1080/09585176.2014.944197

*Mustafa, R. F. (2012). Feedback on the feedback: sociocultural interpretation of Saudi ESL learners' opinions about writing feedback. English Language Teaching, 5, 3–15. doi:https://doi.org/10.5539/elt.v5n3p3

*Nguyen, H. T., & Filipi, A. (2018). Multiple-draft/multiple-party feedback practices in an EFL tertiary writing course: teachers' and students' perspectives. *International Education Studies, 11*. doi:https://doi.org/10.5539/ies.v11n8p1

Nicol, D. (2013). Resituating feedback from the reactive to the proactive. In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education* (pp. 34–49). London: Routledge.

Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: a cognitive model of idea generation in groups. *Personality and Social Psychology Review, 10*, 186–213. https://doi.org/10.1207/s15327957pspr1003_1.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. (2003). Production blocking and idea generation: does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology, 39*, 531–548. https://doi.org/10.1016/S0022-1031(03)00040-4.

Panadero, E. (2016). Is it safe? Social, interpersonal, and human effects of peer assessment: a review and future directions. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of social and human conditions in assessment* (pp. 247–266). New York: Routledge.

Paulus, P. B., & Yang, H. (2000). Idea generation in groups: a basis for creativity in organizations. *Organizational Behavior and Human Decision Processes, 82*, 76–87. https://doi.org/10.1006/obhd.2000.2888.

Peters, M., Godfrey, C., Khalil, H., McInerney, P., Parker, D., & Baldini Soares, C. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare, 13*, 141–146. https://doi.org/10.1097/XEB.0000000000000050.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: a practical guide*. Malden: Blackwell.

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Research Synthesis Methods, 5*, 371–385. https://doi.org/10.1002/jrsm.1123.

*Pitt, E., & Winstone, N. (2018). The impact of anonymous marking on students' perceptions of fairness, feedback and relationships with lecturers. Assessment & Evaluation in Higher Education, 43, 1183–1193. doi:https://doi.org/10.1080/02602938.2018.1437594

Quality Indicators for Learning and Teaching. (2017). 2016 Student experience survey national report. . https://www.qilt.edu.au/docs/default-source/gos-reports/2017/2016-ses-national-report-final.pdf?sfvrsn=14e0e33c_

*Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: mediated by students' perception, moderated by goal orientation. Learning and Instruction, 27, 63–73. doi:https://doi.org/10.1016/j.learninstruc.2013.03.002

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science, 28*, 4–13. https://doi.org/10.1002/bs.3830280103.

Raosoft (2017), Accessed 18 October 2019, retrieved from [online] http://www.raosoft.com

Roseman, I. J., Antoniou, A. A., & Jose, P. E. (1996). Appraisal determinants of emotions: constructing a more accurate and comprehensive theory. *Cognition and Emotion, 10*, 241–277. https://doi.org/10.1080/026999396380240.

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

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

Smith, & Lipnevich. (2018). Instructional feedback: analysis, synthesis, and extrapolation. In A. Lipnevich & J. Smith (Eds.), *The Cambridge handbook of instructional feedback (Cambridge handbooks in*

psychology) (pp. 591–604). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316832134.021.

Smits, M., Boon, J., Sluijsmans, D. M. A., & Van Gog, T. (2008). Content and timing of feedback in a web-based learning environment: effects on learning as a function of prior knowledge. *Interactive Learning Environments, 16*, 183–193. https://doi.org/10.1080/10494820701365952.

Thomson Reuters Endnote X7. (2013). *Clarivate Analytics Endnote (version X7). [Computer software]*. Philadelphia, PA: Clarivate Analytics Available from https://endnote.com/downloads/.

Torrance, H. (2012). Formative assessment at the crossroads: conformative, deformative and transformative assessment. *Oxford Review of Education, 38*, 323–342. https://doi.org/10.1080/03054985.2012.689693.

Van der Kleij, F. M., & Adie, L. E. (2020). Towards effective feedback: an investigation of teachers' and students' perceptions of oral feedback. *Assessment in Education: Principles, Policy & Practice*. Advance online publication. doi:https://doi.org/10.1080/0969594X.2020.1748871.

Van der Kleij, F. M., Adie, L. E., & Cumming, J. J. (2019). A meta-review of the student role in feedback. *International Journal of Educational Research, 98*, 303–323. https://doi.org/10.1016/j.ijer.2019.09.005.

Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: a meta-analysis. *Review of Educational Research, 85*, 475–511. https://doi.org/10.3102/0034654314564881.

Wellington, J. (2010). More than a matter of cognition: an exploration of affective writing problems of post-graduate students and their possible solutions. *Teaching in Higher Education, 15*, 135–150. https://doi.org/10.1080/13562511003619961.

White, H. D. (1994). Scientific communication and literature retrieval. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 41–55). New York, NY: Russell Sage Foundation.

Wiliam, D. (2019). Some reflections on the role of evidence in improving education. *Educational Research and Evaluation, 1*, 127–139. https://doi.org/10.1080/13803611.2019.1617993.

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: a systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*, 17–37. https://doi.org/10.1080/00461520.2016.1207538.