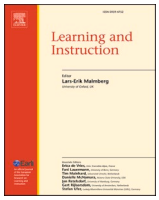
Learning and Instruction 85 (2023) 101745

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc



The eye-mind of processing written feedback: Unraveling how students read and use feedback for revision☆

Renske Bouwer [a,*], Kim Dirkx [b,c]

[a] Institute for Language Sciences, Utrecht University, the Netherlands
[b] Faculty of Educational Sciences, Open Universiteit Heerlen, the Netherlands
[c] Education and Research Department, Zuyd University of Applied Sciences, the Netherlands

ARTICLE INFO

Keywords:
Feedback processing
Eye tracking
Thinking aloud
Revision
Higher education

ABSTRACT

This research aims to promote our understanding of feedback engagement processes in writing tasks using a combination of online and offline measures, including eye-tracking, thinking-aloud, and text-analyses. Study 1 explored how sixteen students read, evaluate, and use feedback for revision. Results revealed three feedback processing strategies: (1) superficial processing ($n = 6$), which is characterized by reading feedback in a linear way, without critically rereading or revising the text, (2) local processing ($n = 6$) in which students switched between reading the comments and the commented text, and (3) deep processing ($n = 4$) in which students integrated the feedback with both commented and uncommented parts of the text and made more substantial revisions. In Study 2, we investigated the local and deep feedback reading strategy in more detail with 41 students using a within-subject design with different types of feedback. Results demonstrated the same strategies among students, but also that the focus of feedback affected students' revision behavior, above and beyond an individual feedback processing strategy. This finding is in line with previous research that emphasized the effects of feedback characteristics on students' use of feedback. By triangulating various process measures, this research is one of the first that provides empirical evidence for different feedback processing strategies among students. These novel insights in individual feedback engagement processing can be used to extend and refine current theories on how, when, and why feedback works and for whom.

1. Introduction

Feedback is one of the most powerful instructional tools for writing (Graham, 2018). Although feedback can be provided in many different modes (e.g., audio or video feedback), in educational practice, most feedback that is provided by teachers on writing products of students is yet predominantly handwritten or electronically written feedback. The importance of these kinds of feedback is evident as it can support students to critically evaluate their text and revise it more effectively (Graham, 2018). Without any feedback, students revise only little, and their revisions are primarily focused on surface-level aspects such as spelling, grammar, and genre conventions, instead of making substantial changes in the content and structure of the text (Faigley & Witte, 1981; Sommers, 1980). Feedback from teachers can direct students' attention from surface-level aspects to the meaning of the text, and hence improve the overall quality of their writing (Paulus, 1999). Moreover, they learn that a text is often not finished at once, but that good writing also involves subprocesses of rereading and rewriting (Hayes et al., 1987). By doing so, feedback can close the gap between current and desired levels of writing performance (Hattie & Timperley, 2007).

Although feedback is seen as an invaluable part of the writing process, it does not necessarily always lead to improved writing performance (Graham, 2018; Hattie & Timperley, 2007; Wisniewski et al., 2020). A critical determinant of feedback effectiveness is the way students actively process the feedback they receive and use it to improve their performance (Boud & Molloy, 2013; Winstone, Nash, Parker, &

☆ These findings were presented in an invited symposium at the online 19th Biennial EARLI Conference for Research on Learning and Instruction in 2021. We have no conflicts of interest to disclose. We thank the teachers, Johan van den Boomen of the Open University, and Louis Lousberg of Delft University (Study 2), to make this research possible and the students Patricia Buijs, Marie-Jose Spek, and Farah Khodabaks for their contribution during the process of data-gathering and analyses. We also thank Desirée Joosten - ten Brinke and Naomi Winstone for their insightful comments on a previous version of this paper.

* Corresponding author. Institute for Language Sciences, Utrecht University, Trans 10, 3512, JK, Utrecht, the Netherlands.
E-mail address: r.bouwer@uu.nl (R. Bouwer).

https://doi.org/10.1016/j.learninstruc.2023.101745
Received 16 December 2021; Received in revised form 28 November 2022; Accepted 1 February 2023
Available online 8 February 2023

Rowntree, 2017). To revise a text successfully in response to feedback, students should "at least read and take seriously those comments" (Ferris, 1995, p. 48). In addition, for effective revisions, students need to critically evaluate and use the feedback by linking the comments to their own text and deciding what to change in the text and how (for a cognitive process model of revision, see Hayes et al., 1987).

But to what extent do students actively and critically engage with the feedback they receive? So far, empirical studies on feedback use in writing have predominantly focused on changes in the final product instead of the underlying processes, showing that students use teacher feedback only to a limited extent (Hyland, 1998). In addition, studies that focused on feedback processes relied only on self-reported data (Jonsson & Panadero, 2018). Therefore, studies revealing the online processes of feedback engagement are still underrepresented. It has been suggested, however, that during feedback processing students may experience psychological barriers that inhibit them from using feedback. Based on retrospective focus group interviews, Winstone, Nash, Rowntree, and Parker (2017) revealed four of those barriers, which are related to students' lack of awareness, cognizance, agency, and volition in using (effective) strategies to take action in response to feedback. These barriers suggest that students may engage with feedback quite differently. This can be confirmed by the work of Burke (2009), in which students indicated that they received inadequate guidance on what to do with the feedback. Additionally, Sutton (2012) argued that teachers use complex academic language in their feedback which makes it hard for students to adequately understand and use the feedback. To design more effective feedback interventions, it is needed to further unravel the strategies of how students process and use teacher feedback to improve their writing.

In this research, we present the results of two studies that aim to reveal the underlying processes of feedback engagement in writing. To analyze how students process feedback in fine-grained detail, we combined several online and offline process measures, such as eye tracking, thinking aloud, keystroke logging, and text analyses. Together, this research will extend current theories on how students actively engage with teacher feedback in the context of writing. It also provides a new methodological perspective to study the underlying processes of feedback engagement.

1.1. Underlying processes of feedback engagement

An important determinant in how actively and critically students engage with feedback is mindful cognitive processing, which refers – according to Bolzer et al. (2015) – to "how deeply the (peer)feedback has been cognitively processed and understood" (p. 425). This involves three different (meta)cognitive strategies. First, students need to read the feedback and relate it to their work (Berndt et al., 2018; Bolzer et al., 2015). Second, they need to evaluate the quality and usefulness of the feedback by comparing it to what their current understanding and performance, and deciding if and how they want to use it to improve their work (i.e., making evaluative judgments about the feedback; Boud et al., 2018; Carless & Boud, 2018; Nicol, 2020). Finally, they need to use the feedback to revise their work (i.e., taking action; Carless & Boud, 2018; Winstone, Nash, Parker, & Rowntree, 2017). The first two strategies in which students make sense of feedback information are considered to be internal processes (De Kleijn, 2021). The third strategy, in which students act upon the feedback, may involve both internal and external processes (De Kleijn, 2021; Molloy et al., 2020). That is, feedback can be used by students to improve their text, but also their learning and understanding, for instance by changing mental concepts or increased feedback recall, without this being directly visible by external revisions in the text (Bolzer et al., 2015). Online measures can provide important insights into the extent to which students are involved in the underlying and internal processes of reading, evaluating, and using feedback. Below, we will describe the different feedback processes in more detail and propose ways to measure them in real-time.

Reading the feedback. A crucial prerequisite of effective feedback is that students "at least read [and take seriously] those comments" (Ferris, 1995, p. 48). However, there is rather little research which investigates to which extent students read the feedback they receive and the research that has been done has led to different results. Research of Cohen (1987), for example, has shown that students do not always read the feedback they receive, especially when the assignment is already graded. Cohen and Cavalcanti (1990) on the other hand showed that students carefully read all the feedback they received. However, a major problem with previous studies is that they are mainly based on self-reported data which requires students to report retrospectively how they have read the feedback. A drawback of self-report methods is that they ask students to report how they evaluate the feedback after some time has passed. Thus, they tap into long-term memory information which can give a distorted picture of reality.

A more realistic tool for measuring reading behavior is eye-tracking. Eye-tracking provides insight into where a person is looking, for how long, and in what order (Holmqvist et al., 2011). Since eye movements are closely linked to the focus of attention when students process information (Just & Carpenter, 1980), one gets insight into the "online" cognitive processing of - for example - reading a text (Hyönä et al., 2003; Just & Carpenter, 1980). Although eye tracking is a widely and long-used method in reading research (see Rayner, 1998 for an overview), and recently also in writing research (Wengelin et al., 2019), there is only a small number of studies that have used it to analyze feedback engagement processes (Alqassab et al., 2018; Berndt et al., 2018; Bolzer et al., 2015; Paulson et al., 2007). For example, Bolzer et al. (2015) used eye tracking to investigate the cognitive processing of peer feedback. In this study, eye fixation duration and number of transitions between feedback and text were used to infer how deeply and mindfully students processed the peer feedback. The results showed that when no justifications for the feedback were provided, students invested more time and effort in reading the feedback and integrating the feedback more deeply with the text. Longer reading times and increased integrative effort were positively correlated and they were both associated with improved feedback recall. Research of Berndt et al. (2018) showed that glance durations (i.e., the sum of all fixations and saccades between and within the area of interest (AOI), including the duration of the initial saccade entering the AOI) were negatively correlated to revision behavior, which was mediated by the amount of information presented in the feedback. This suggests that students should not be "overloaded" with feedback, as the number of feedback comments increases processing time but at the same time also decreases the probability of revision. Together, these findings show that eye tracking data based on fixation durations and gaze transitions can provide in-depth information on how students read teacher feedback comments and link it to their own text.

Evaluating the feedback. Next to reading and understanding the feedback by linking it to the text, students need to evaluate its quality and usefulness. Depending on these evaluative judgments, students decide what actions they will take (De Kleijn, 2021). Therefore, revealing the underlying process of how students evaluate the feedback can reveal interesting information on why they are willing to use the feedback or not. Yet, most of the research that has been conducted on students' judgments of feedback used self-report surveys or retrospective interviews (Winstone, Nash, Parker, & Rowntree, 2017). A second limitation is that in most research students' feedback judgments have not been related to students' reading behavior and/or subsequent actions.

Capturing "mental notes" might be a more valid way of measuring students' judgments as they more directly elicit students' feedback evaluations, that is, immediately when these evaluations are taking place. As such, mental notes represent the information processed in working memory (Van Gog et al., 2009). This can provide information on students' instant evaluative comments on the feedback (Berndt et al., 2018). A methodology that is appropriate to measure mental notes is

thinking-aloud (Ericsson & Simon, 1993), in which students verbalize what they think when performing a task. Thinking-aloud data can also help to interpret data that is collected via eye-tracking (Van Gog et al., 2005), by revealing why a student is looking at certain information and neglecting other sources of information. Thus, triangulated with eye-tracking, think-aloud data can help to further unravel the way in which students process feedback.

Acting upon the feedback. After evaluating the feedback, students need to use the feedback for revising their text. Revision can be defined as any change at any moment during writing (Fitzgerald, 1987). It involves identifying differences between the intended and the written text, deciding what could or should be changed, and how the desired changes should be made, and then ultimately also make those changes in the text (Fitzgerald, 1987; Hayes et al., 1987). Thus, revision processes are partly internal, in which changes are made in the writer's mind, and external, in which changes are visible in the written product.

Revisions in the text can be quite objectively analyzed by text-analysis. In this analyzing technique, the old and new text are compared by researchers, either by hand or automatically by software programs such as Inputlog (Leijten & Van Waes, 2013), in which all keystrokes and deletions are automatically registered. Text-revisions can be categorized as additions, deletions, or changes in the existing text, and they may affect the meaning of the text or not (Faigley & Witte, 1981). Previous studies using text-analysis show that students use between 49% and 90% of the feedback for revisions (Beason, 1993; Ferris, 1995; Paulus, 1999; Treglia, 2009).

However, text-analyses do not reveal the changes that are made in the writer's mind. These mental revisions provide, however, very valuable information about the way students respond to feedback. Albeit most students seem quite happy to get feedback, not all students seem to possess fruitful strategies for dealing constructively with it. For instance, previous research showed that many students use the feedback they receive rather passively (e.g., making mental notes; Pokorny & Pickford, 2010; Williams & Kane, 2009), or not at all (e.g., by erasing the problematic issues raised by the teacher; Hyland, 1998). Furthermore, McCurdy (1992) and Ferris (1995) showed that students may also turn their attention to the teacher, peers or their textbooks for further guidance or instructions. Triangulating text-analyses with think-aloud data, can help to also capture these mental revisions and hence to unravel different strategies in feedback use.

1.2. Research aims

In this research we aim to extend current theories of feedback engagement by unraveling the underlying processes of feedback engagement. In two studies we triangulate online process measures of eye tracking, thinking aloud, and keystroke logging to capture the internal feedback processes, and combine them with text analyses to also measure external feedback processes. In addition, by combining thinking-aloud protocols with more unobtrusive and automatic process measures such as eye tracking, we diminish the risk of biased results due to socially desirable verbalizations. The first study is an exploratory case-study in which we explore in an authentic setting how premaster students read, evaluate, and use teacher feedback for revision. The aim of the second study is to replicate these findings in a different context, as well as to further examine the effects of different modes of teacher feedback on how students process and use feedback for revision. Together, these studies empirically investigate in fine-grained detail how students engage with teacher feedback and revise accordingly.

2. Study 1

This case-study explored in an authentic setting how premaster students process and use teacher feedback for revision. Specifically, we investigated the following research questions.

1. How do students read teacher feedback?
2. How do students evaluate teacher feedback?
3. How do students revise in response to feedback?
4. To what extent can we distinguish between individual feedback processing strategies, based on the relations between how students read, evaluate, and use feedback for revision?

2.1. Method

Participants. Twenty-one Dutch-native speaking university students agreed to participate voluntarily in this study. They received a voucher of 15 euros after participation to compensate for travel expenses. The students followed a one-year pre-master's program on Management Science. Three students did not show up for the data collection and for two students the eye movement data could not be used due to equipment failure or problems with eye tracking calibration. Therefore, the final sample consisted of sixteen participants: four females and twelve males. They were aged between 22 and 40 years.

Materials and Procedure. During the course, students received authentic teacher feedback on their written reports. This feedback was directly embedded in the text (see Fig. 1). The reports were about three pages long and structured by five leading questions in which they had to interpret statistical results of a reliability analysis in SPSS for a questionnaire. Feedback was provided by one of the two teachers of the course, who had over 10 years of teaching experience. The content of the feedback was based on a collection of exemplary comments gathered over the years. On average, teachers provided ten feedback comments per student ($SD = 2.61$; min. = 7, max. = 19), with 166 comments in total. The report was not graded, and students were instructed to revise their report based on the feedback as a preparation for the summative assessment two weeks later.

To explore how feedback is read, students were asked by the researcher during class-time to participate in the study. When students agreed upon participation, they were contacted by the researcher and an appointment was made. When the participant arrived for participation in the study, they were instructed as follows: *"Welcome! Please take a seat. In a moment you will see on this computer your teacher's feedback on the assignment you turned in to your teacher last week. We want you to work on the assignment as you normally would when you receive feedback to improve the assignment at home. We ask you to think out loud while you are doing this*".

After the instructions, the feedback was shown on a remote SMI 250HZ eye-tracker which was fixated under a computer screen and sampled data at the rate of 250Hz. Students were seated about 60 cm from the screen. Their text, including the feedback, was displayed as a MSWord© file on the screen. Students were instructed to read the feedback and respond to it as they would normally do while thinking aloud and making revisions in the text. We used screen recordings with SMI-software to track students' eye movements while they were scrolling through the document. This was preceded by a five-point calibration with black crosses on a plain, white background.

For the think-aloud procedure a brief instruction page was used as training material. This included the following specific instructions: "*While you are working on the text, please think out loud, just as if you were speaking to yourself. Act as if you are alone in the room speaking to yourself. If you are silent for any length of time, I will remind you to keep talking aloud. Do you understand what I want you to do?*" (Ericsson & Simon, 1993, p. 376). Think-aloud data was recorded using the microphone of a Logitec webcam. Students were able to make revisions in the text immediately. After 30 min, students were asked if they were ready. When they finished, they were thanked for their participation and received the 15 euros. Students who did not volunteer to participate, received their feedback via e-mail by the teacher.

Data analysis. First, we analyzed students' reading behavior by categorizing their gaze patterns into two subprocesses: (1) linear or

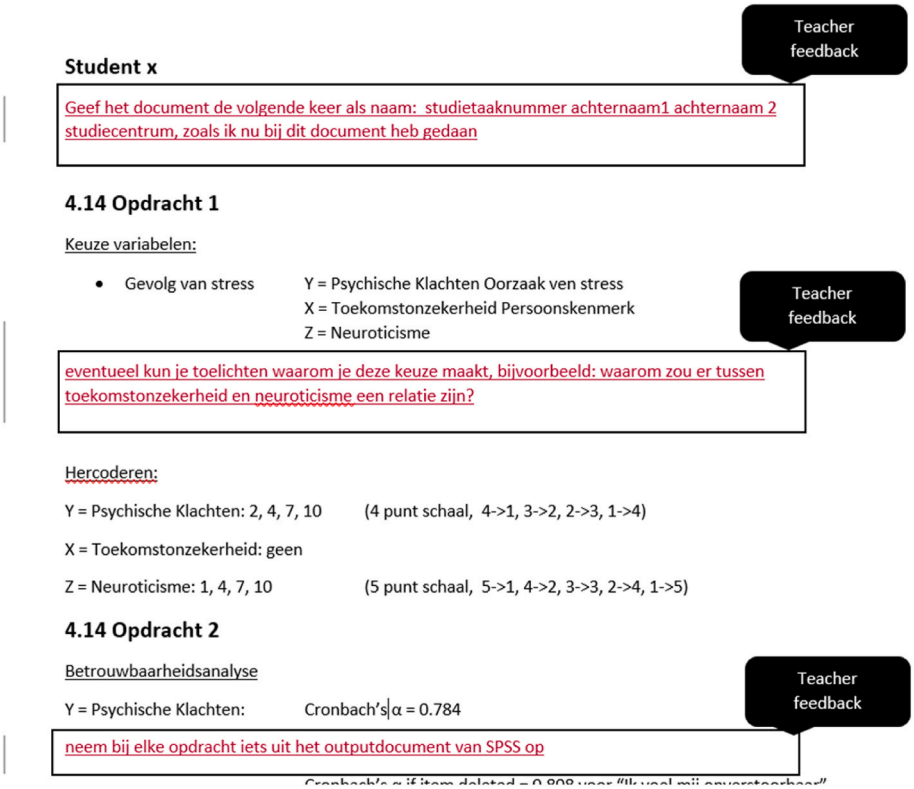


Fig. 1. Example of a student's work with embedded teacher feedback (in Dutch).

cyclic reading of the feedback, as indicated by reading feedback comments one by one or by going back and forth between different feedback comments, and (2) local or global integration of feedback and text, as indicated respectively by transitions between feedback and specific commented parts of the text or transitions between feedback and larger parts of the (uncommented) text. Based on the combination of these two subprocesses, we can infer how deep students read the feedback, ranging from passive and superficial reading (i.e., linear reading of the comments and local integration with the text) to active and deep reading (i.e., cyclic reading and global integration with the text). There was a substantial level of agreement between the two coders (first and second author), with a Cohen's kappa of .71.

Second, we transcribed and coded all verbal reactions to the feedback comments as measured by the thinking aloud protocols. The first and second author decided together whether each reaction can be considered as an evaluation or a (mental) revision. The evaluations and mental revisions were then coded separately using a bottom-up coding process in which both authors coded 20% of the transcriptions individually and discussed the discrepancies in the coding. A second round of coding resulted in a satisfying interrater agreement for both evaluations ($\kappa = 0.86$) and mental revisions ($\kappa = 0.72$). The second author then coded the rest of the evaluations.

Third, for each feedback comment we indicated whether the student made a revision in the text by comparing the old and new documents. For each revision, we also coded whether it was a deletion, addition, or change in the text (cf. Faigley & Witte, 1981).

Fourth, we qualitatively described individual feedback processing strategies by triangulating the eye-tracking data, thinking-aloud protocols, and text analyses, in order to gain in-depth insight in how students read the feedback and try to understand and evaluate the feedback by linking it to their own text. These descriptions also included information about whether the student looked at all feedback comments, made transitions between the text and feedback, and made revisions in the text, either directly, after reading all comments or not at all. These qualitative descriptions were made by the second author and discussed and complemented by the first author.

2.2. Results

Students engaged on average 15 min ($SD = 12.03$) with the assignment. There were, however, large differences between students, ranging from 4.42 min to 52.50 min. Below, we will further unravel these individual differences. We will first describe the sub processes of reading, evaluating, and using the feedback, after which we will integrate these subprocesses into individual feedback processing strategies.

Reading the feedback. The gaze patterns revealed that all students read all the feedback in this time, but they differed in how deeply they read the feedback and whether they integrated the comments to their text. Four students showed a superficial reading process, by reading the comments only once and in a linear way from the top to the bottom of the document without looking back to any of the previous comments. These students also used a local integration approach, by looking only at the commented parts of the text. The other twelve students read the feedback in a cyclic manner by moving back and forth between comments. Half of them linked the feedback only to commented parts of the text (i.e., local integration process; $n = 6$), while the other half also looked at parts of the text that received no specific comments (i.e., global integration process; $n = 6$).

Evaluating the feedback. The thinking aloud data revealed that

Table 1
Frequencies and percentages for categories of feedback evaluations.

Category	N	%	Nr of students
Cost-effort estimation	17	10.2	8
Recognizing the feedback	14	8.4	5
Understanding the feedback	8	4.8	5
Valuing the feedback	5	3.0	3
No evaluation	111	66.9	16

while students read all the feedback, they hardly explicitly evaluated the comments. Table 1 shows that students evaluated only one-third of all comments. These evaluations were mostly cost-effort estimations (10%), in which students ($n = 8$) evaluated the importance of the feedback and estimated whether it was worth the effort to act upon the feedback by revising the text. For instance, student 3 said "*the teacher says 'maybe', so it is not that important*". Five students indicated to recognize some of the feedback they received by the teacher (8% of the comments). For example, student 11 said "*Yes, I totally agree*" and "*that is true, we did not describe this in the conclusion*". There were also some instances (5% of the comments) in which five students indicated whether they understand the feedback or not. For instance, student 4 said "*ok, I assume he [the teacher] means that I mention the items it concerns*", or "*it is not clear to me what he means with this*". Three students indicated for five comments (3%) whether they valued the feedback or not. Student 12 valued the feedback in a positive way by saying *"I really like the tips in the feedback that I received from the teacher"*.

Using the feedback. The thinking aloud protocols and text analysis revealed that students acted upon 37% of the feedback comments, see Table 2 for an overview of the type of actions in response to the feedback. Most of the revisions in the text were additions (10%), followed by changes in the text (2%) and deletions (3%). Eleven students also made mental notes for future assignments (17% of the comments). For instance, student 5 said "*I see that I could better use my own words. This is maybe something for 4.15 [the next assignment]"*. For some comments, students indicated that they needed external information, either from the teacher, classmates, or source materials. For example, student 1 stated *"I first need to check the correlation coefficient in the output file"*.

Aggregation at the individual student level showed that the percentage of revisions per student ranged from 0 to 89%. Nine students revised their text directly, by making additions ($n = 7$), deletions ($n = 3$), and/or changes ($n = 4$). Eleven students decided to postpone all or some of their revisions until later. Two students hardly acted upon the feedback and three students did not act at all, neither by revisions in the text nor by making mental notes for future assignments.

Feedback processing strategies. By combining the subprocesses of how students read, evaluate, and use feedback, we were able to determine three individual feedback processing strategies.

1. Superficial feedback processing strategy ($n = 6$): students who used this strategy read feedback in a linear way, without critically rereading comments and by only locally integrating the feedback with the text. They hardly made any critical evaluation of the feedback, nor did they make substantial revisions in the text. If they did make revisions, they did so directly and at a rather surface level of the text. For instance, student 11 received the feedback "*Strictly speaking, you write that omitting 10 items is sufficient, but in the sentence above I cannot deduce whether you have included those 10 items in the variable or not. Formulate this more precisely, preferably also by a formula"*. The student responded to this feedback by stating "*you're absolutely right*", and by adding only the formula but without reformulating the sentence. The student did not take any action in response to the other comments.
2. Local feedback processing strategy ($n = 6$): local feedback engagers are characterized by a cyclic but local reading behavior. They started at the top and first went through all the comments. They only looked back at the commented parts of the text and paid little or no attention at all to parts that received no specific comment. In some cases, students already started revising, but most revisions were performed after all feedback was read. Students with this processing strategy made more critical evaluations and revisions in response to the feedback than students with a superficial processing strategy, however, most revisions were still at a surface level of the text. A main reason for these students to not revise in response to feedback or postpone their revisions is that they did not fully understand the feedback. There were also students who made cost-effort estimations and decided that the comment was not important or worth the effort to revise.
3. Deep feedback processing strategy ($n = 4$): students who used this strategy also read the feedback more than once (i.e., cyclic approach), but they put more effort in integrating the comments to the text by reading the commented parts of the text as well as the parts that received no specific comments. This suggests that they were trying to understand the feedback of the teacher more deeply. In addition, they evaluated the feedback more critically and used it more substantially, also for meaning-level revisions throughout the whole text. For instance, student 9 responded to one comment by adding new text, deleting tables, and restructuring information.

Table 2
Frequencies and percentages for categories of feedback use.

Category	N	%	Nr of students
Additions in the text	17	10.2	7
Deletions in the text	4	2.4	3
Changes in the text	5	3.0	4
Mental note for future	28	16.9	11
Ask for help/look at materials	7	4.2	2
No action	105	63.3	3

3. Study 2

The primary aim of this follow-up study was to validate the local and deep feedback processing strategies as found in Study 1. A second aim was to empirically investigate how these processing strategies affect revision behavior, above and beyond specific feedback characteristics, such as the focus and directiveness of the feedback. The relationship between individual processing strategies and feedback characteristics has not yet been systematically investigated in previous research (Jonsson & Panadero, 2018, p. 541). To do so, we adopted a within-subject design in which we systematically varied the focus of feedback comments, i.e., feedback focused on surface-level or meaning-level issues in the text, and the directiveness of the feedback, i.e., track-changes or directive comments in the text or questions and facilitative suggestions next to the text. Specifically, we investigated the following research questions.

1. To what extent can the differences between a local and deep feedback processing strategy, as determined in Study 1, be replicated in a different sample with students from a different educational context?
2. What are the effects of the focus and directiveness of the feedback on students' reading and revision behavior?

3.1. Method

Participants. In total, 47 third-year Bachelor students in Architecture, Urbanism, and Building Sciences at the Technical University of Delft, the Netherlands participated in this study. This was 44% of all third-year students in this Bachelor (n = 106) who were invited to participate on a voluntary basis. For three students, the eye movement data could not be used due to equipment failure. Three students were excluded from the data-analysis as they were English and experienced problems in reading and revising the text, which was written in Dutch. The final sample therefore consisted of 41 Dutch-native speaking students: 24 male, 16 female, and one student didn't indicate a gender. Their mean age was 22.3 years (SD = 1.75), ranging from 20 to 27 years.

Design and power. We used a scenario-based design in which all students were presented with the same text of average quality written by a student in a previous year. The text included ten feedback comments based on authentic teacher feedback which varied in focus (i.e., higher versus lower order feedback) and directiveness (i.e., facilitative

questions and suggestions or directive comments). The scenario-based approach enables a systematic and controlled investigation of the effects of feedback characteristics on feedback processing strategies (Berndt et al., 2018). Prior research has shown that scenarios invoke almost identical reactions from persons, compared to real situations (Robinson & Clore, 2001).

A statistical power analysis for multilevel logistic regression shows that, with a two-tailed test with a binomial distribution and alpha of .05, a probability of .30 for revision in response to directive feedback and 0.48 for facilitative feedback (Khodabaks, 2020), and feedback comments that are nested within students resulting in a design effect of 1.45 (cf. Hox et al., 2018), there is an 89% change of correctly rejecting the null hypothesis that the probability of revision is not associated with the modality of feedback. According to Hox et al. (2018), this can be considered as a high power. With the relatively small sample of students, we can also achieve sufficient statistical power to disentangle between the two feedback processing strategies based on k-means cluster analysis (Dalmeijer et al., 2022).

Ethics. This study is part of research on feedback and assessment within the former Welten Institute (now the Faculty of Educational Sciences of the Open University in the Netherlands), for which ethical approval has been granted by the committee Ethical Review Research (cETO) of the Open University. Two weeks before the start of the research, all participants were informed orally by their teacher about the purpose, relevance, procedure, and duration of the research. They were instructed that participation in this study was on a voluntary and anonymous basis, and that individual results would not be shared with the teacher. Students were able to ask clarifying questions to both the teacher and researchers. One week before the start of the study, all students received an email from the researchers with further information about when and where they were expected. One day in advance, the same email was sent again as a reminder. At the start of the study, the subjects had to sign a declaration of consent for participation. They also gave their consent for the storage of the collected data for 10 years on protected servers of the Open University.

Materials and procedure. Students were asked to participate via e-mail by the teachers. Participation was voluntary and students did not receive any incentive or reward for participation. All participants received the same text. The text was of average quality and carefully selected from a sample of 21 authentic student papers that were written in the previous year in a course on academic design reflections. In this course, students have to write three reflective essays on their design process using scientific literature. The text was selected by the two main researchers and two research assistants who evaluated the quality of texts using comparative judgment (Verhavert et al., 2019). We selected one of the texts that was ranked in the middle (i.e., a logit score around zero) and for which the rater disagreement was lowest (logit = −0.25, $SE = 0.61$).

The feedback on the average text consisted of ten comments based on authentic teacher feedback for comparable writing products. The focus and directiveness of the feedback varied within participants. Four comments were focused on lower order aspects, such as grammatical errors, language use, and references in the text. This lower order feedback was provided to students either as directive comments or by track-changes in the text. The other six feedback comments were focused on higher order aspects in the text, such as the content and structure of the information (e.g., unclear research question and definition). Three of these comments were presented as directive feedback next to the text (e.g., *'remove the personal reflection here'*) and three comments as facilitative feedback (e.g., '*the five elements of Van Doorn* et al. *are always present in a design process, what does that imply for your essay*?').

In the first phase of the study, participants were presented with the text and the feedback on a computer screen and instructed to carefully read it as they were given time to revise the text in the second phase. To measure students' attention to the feedback and the accompanying text during reading we registered their eye movements using a remote desktop eye tracking system SMI with 250 hz sampling rate that was connected to the computer screen. Before presenting the feedback, students' eye movements were calibrated. In the second phase, participants were replaced from the computer screen to a laptop, where they were requested to revise the text as if it were their own. This setup allowed us to capture a better quality of eye tracking data, as students did not work on their texts while eye movements were tracked. Students' revision behavior was automatically tracked by the keystroke logging program *Inputlog*. This program registers all keystrokes, mouse clicks and pausing times during the revision process.

Data analysis. For the analysis of the eye-tracking data, we used software Begaze 3.7. As in study 1, we determined students' feedback processing strategy based on how deeply students read feedback and integrate it to the text. In this study, however, we were only able to distinguish between a local and deep feedback processing strategy, as students were not able to revise immediately after reading a feedback comment as they would do in a superficial feedback processing strategy. To determine students' processing strategy based on the eye-tracking data, we first defined three different types of areas of interest (AIOs): feedback comments, commented parts of the text, and uncommented parts of the text. Then, we used the normalized dwell time and the number of revisits for the three predefined AOIs to reveal the extent to which students read and switched between feedback comments (i.e., linear or cyclic reading of feedback) and between feedback and commented or uncommented parts of the text (i.e., local or global integration). To validate the robustness of the two feedback processing strategies, we conducted *k*-means cluster analysis based on the normalized dwell time and the number of revisits for the three predefined AOIs.

Revisions were categorized as either meaning or surface changes (Faigley & Witte, 1981). Inter-rater agreement for two coders was high ($\kappa = .91$). To analyze the probability of revision in response to different types of feedback we used multilevel logistic regression modeling with revision as a categorical outcome variable. In this multilevel analysis we accounted for variance within students (Level 1) and between students (Level 2). In the fixed part of the model we estimated an intercept, indicating the average probability of revision, given a random feedback comment and a random student. We also added the two feedback characteristics (i.e., focus and directiveness) as fixed variables, in order to analyze whether students were more likely to revise in response to particular types of feedback. Furthermore, individual feedback processing strategy is added as a fixed variable, to analyze whether the extent to which students engage with the feedback increases the probability of revision. The parameter estimates are presented in logits, which are a nonlinear transformation of the probabilities. To enhance interpretation in the text, the logits are transformed back to probabilities of occurrence.

3.2. Results

The eye tracking data reveal that students vary considerably in how deeply and actively they engage with the feedback, as reflected by differences in the way they read the feedback and integrate it with their text. A visual inspection of the gaze data suggests that, as in Study 1, feedback was either read in a linear (only once) or cyclic manner (going back and forth between comments). We used a *k*-means cluster analysis to distinguish between the local and deep feedback processing strategy, based on normalized dwell time and revisits of the commented and uncommented parts of the text as well as revisits of the feedback comments. The two-cluster solution seemed to fit the data perfectly as stable convergence of the two clusters was already achieved after three iterations. Also, the number of cases for each cluster was adequate, respectively 14 and 27. Adding more clusters to the analysis resulted in clusters involving only one student. Fig. 2 shows how the eye tracking variables cluster around the two processing strategies of the two-cluster solution.

Table 3 demonstrates that students with a deep feedback processing

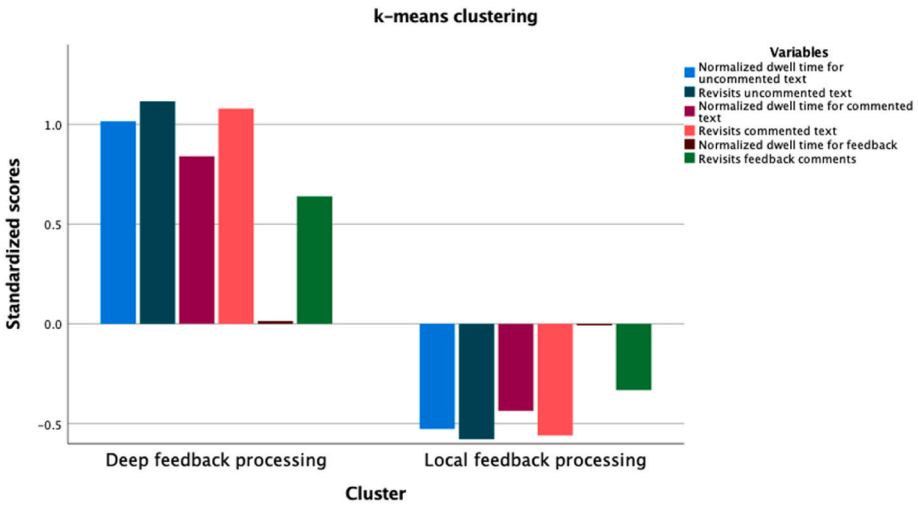


Fig. 2. Standardized score means for the cluster variables of the two feedback processing strategies based on *k*-means clustering.

Table 3
Means, standard deviations, and analyses of variance for the cluster variables.

	Deep feedback processing (n = 14)		Local feedback processing (n = 27)		F (1, 29)	η^2
	M	*SD*	*M*	*SD*		
Normalized dwell time (s) for uncommented text	114.14	32.57	48.62	26.93	47.28***	.55
Revisits uncommented text	10.52	3.39	3.44	1.83	76.11***	.62
Normalized dwell time (s) for commented text	187.37	49.18	113.50	44.78	23.47***	.38
Revisits commented text	8.30	2.37	3.66	1.38	63.34***	.59
Normalized dwell time (s) for feedback comments	71.40	57.33	70.24	50.00	0.004	.00
Revisits feedback comments	3.60	2.38	1.87	1.00	10.84***	.22

*$p < .05$. **$p < .01$. ***$p < .001$.

strategy (cluster 1; $n = 14$) spent significantly more time reading both the commented and uncommented text than students with a local feedback processing strategy (cluster 2; $n = 27$; $F > 47.28$, $p < .001$). They also made more transitions between the text and the feedback as reflected by the number of revisits, indicating that they integrated the feedback more deeply to the text. There was no difference between the processing strategies regarding the duration for reading the feedback comments ($F(1, 29) = 0.004$, $p = .95$). Fig. 3 illustrates the gaze behavior for a local and deep feedback processing strategy.

Results from the multilevel logistic regression model (see Table 4) shows that students were very likely to make revisions in the text in response to feedback, with an average probability of revision of 96% (Logit = 3.16, $SE = 0.50$; $t = 6.37$, $p < .001$). Revisions consisted for 54% of surface changes. The other 46% were meaning-level changes, with most changes at the sentence or paragraph level (41%) and only 5% macro-structural changes that affected the overall meaning of the text.

Regarding the effects of feedback characteristics, it was shown that only the focus of feedback had a significant effect on the probability of revision, $F(1, 406) = 20.16$, $p < .001$. Specifically, feedback on meaning-level aspects decreased the probability of revision with 12% in comparison to surface feedback (Logit = −2.02, $SE = 0.45$, $t = -4.45$, $p < .001$). There was no effect of the directiveness of the feedback on revision, $F(1, 406) = 0.02$, $p = .88$. There was also no main effect of individual feedback processing strategy on the likelihood of revision ($F(1,$

Fig. 3. Illustration of gaze behavior for a student with a local (left) and deep feedback processing strategy (right).

Table 4
Parameter estimates for planned multilevel model of the probability of revision (N = 410).

	Logits (SE)
Fixed effects	
Intercept	3.16*** (0.50)
Focus of feedback	−2.02*** (0.45)
Directiveness of feedback	−0.05 (0.31)
Feedback processing strategy	0.17 (0.42)
Random effects	
Students (S_s^2)	.76* (.37)
Goodness of fit	
AIC	2066.08
BIC	2070.08
-2Loglikelihood	2064.07

Note. Standard errors are in parentheses. All *p*-values in this table are two-tailed.
*$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

406) = 0.17, $p = .68$), nor was there an interaction effect between feedback processing strategy and the focus of feedback ($F(1, 406) = 1.59$, $p = .21$).

Even after controlling for the effects of feedback focus, there was large variability in revision behavior between students ($S^2 = 0.76$, *Wald* $z = 2.06$, $p = .02$), with a 95% confidence interval for the probability of revision ranging from 0.57 to 0.87.

4. Discussion

In this research, we set a first step to empirically investigate how actively and deeply students engage with teacher-provided feedback on a formative writing task and how they use it for revision using a triangulation of online and offline measures. In two studies, we demonstrated that students process feedback in different ways. In particular, the results of Study 1 revealed three different feedback processing strategies: (1) a superficial processing strategy, which is characterized by students who read feedback only once and in a linear way, from top to bottom, without (critically) rereading the feedback or the text, (2) a local processing strategy, in which students read the feedback more than once and regularly switch between the comments and the commented text, and (3) a deep processing strategy, in which students integrated feedback with both the commented and uncommented parts of the text. Furthermore, the combined results of the qualitative data suggests that the depth by which students were going back and forth between feedback comments and integrated the feedback to their own text is associated with a more critical understanding and evaluation of teacher feedback and more substantial revisions throughout the whole text. This finding is in line with previous research of Bolzer et al. (2015), showing that deep cognitive processing of feedback is associated with more effective feedback use.

In Study 2, we were able to replicate the local and deep feedback processing strategy in a different context and with other students. We investigated the differences between these two active processing strategies in a more systematic and detailed manner and included characteristics of the feedback message as additional variables. To do so, we performed a cluster analysis on reading times and number of revisits for feedback comments and the commented and uncommented text and we related the processing strategy to revision behavior. In line with Study 1, the results demonstrated that deep feedback processing was associated with longer reading times and more revisits for both the commented and uncommented parts of the text than a local feedback processing strategy. This indicates that students with a deep feedback processing strategy generally put more effort in integrating the feedback with the whole text. Remarkably, this was not related to more revisions in the text: students with a deep feedback processing strategy were as likely to revise in response to feedback as students with a local feedback processing strategy. Moreover, revision was affected by the focus of feedback. In general, students were more likely to revise in response to surface-level feedback than to meaning-level feedback, regardless of the depth with which they processed the feedback. This finding demonstrates how the content of feedback can promote students' proactive use of feedback for revision (Winstone, Nash, Parker, & Rowntree, 2017). The directiveness of the feedback did, however, not affect students' revision behavior.

That there was no relation between the feedback processing strategy and feedback use in the second study might be explained by the overall high percentage of revision. Whereas students in the first study only acted upon 37% of the feedback, students in the second study revised in response to 85% of the feedback comments. This increased probability of revision could be due to the study's design. Students did not revise their own work, but a text of average quality that was written by a peer. This scenario-based approach enables a systematic and controlled investigation of the effects of feedback characteristics on feedback processing strategies (Berndt et al., 2018). Prior research has shown that scenarios invoke almost identical reactions from persons, compared to real situations (Robinson & Clore, 2001). However, work of Elving (2019) suggests that revising work from others is easier than revising one's own work, which might also lead to more revisions in response to feedback. Furthermore, in Study 2, students read the text and the feedback before they were able to revise the text. Although this design gave us the opportunity to separate the feedback reading behavior from the revision behavior, it might have also given students an advantage by limiting the cognitive capacity that is needed for (1) understanding and evaluating the feedback and (2) acting upon it, resulting in relatively more revisions (Hayes et al., 1987). Further intervention research is needed to investigate whether students' feedback engagement processes can be improved by teaching them to apply deep processing strategies to example texts before they learn to apply these strategies to their own work, as well as by separating the processes of reading and evaluating feedback from using it for revision. Further research is also needed to experimentally investigate the relationship between a deep feedback processing strategy and the quality of the revisions in more detail, and to test whether these findings are also also found using different study designs to resolve some possible limitations due to the current design (e.g., small sample sizes, observation bias, and selection effects).

The current findings can be placed in a more general discussion evolving around the concept of feedback literacy (De Kleijn, 2021; Molloy et al., 2020). According to Sutton (2012), feedback literacy is the ability to read, interpret, and use written feedback. Sutton concludes that "for some learners, reading, interpreting and acting upon feedback is a significant challenge" (p. 38). The academic language that teachers use when providing feedback seems one such barrier to interpret and use feedback effectively. As a result, students may develop a poor conceptualization of feedback, which is illustrated by one of the students in Sutton's article: "you learn what they want and then you incorporate that into the next piece of work" (p. 38). This rather passive evaluation of feedback was also demonstrated in the first study. Students either just took over the comments of the teacher immediately without any critical evaluation of the feedback or they made cost-effort estimations in which they explicitly decided whether it is worth the effort to revise for getting a good grade on the final exam. Both these responses can be considered as a rather superficial and passive way of feedback engagement (Jonsson & Panadero, 2018; Winstone, Nash, Parker, & Rowntree, 2017). However, the results also demonstrated a deep level feedback processing strategy among students, indicating that there are students that are eager to learn from the feedback.

Another limitation of the current study is that we measured students' evaluation processes only in the first study, but not in the second study. The reason that we did not continue to measure students' evaluation of the feedback has to do with concerns regarding the think-aloud method

that we used which might have caused some unwished effects such as cognitive overload or observer effects (e.g., the Hawthorne effect, see Sommer, 1968; or McCambridge et al., 2014). For instance, the number of verbalizations in the first study was rather low, which might suggest that reading and interpreting feedback is already so cognitively demanding that, for most students, it is hard to also verbalize one's thoughts. More frequent reminders to verbalize the thinking processes can potentially disrupt and bias the reading and revising behavior of students, also known as the observer effect (Ericsson & Simon, 1993; Sommer, 1968). Therefore, we decided in the second study to use only automatic process measures that did not affect the reading and revising processes. However, by doing so, we cannot explain why students in this study were more likely to revise in response to surface-level feedback instead of meaning-level feedback. Future research is needed to investigate students' reasons for using feedback (or not) in more detail, for instance by using cued retrospective recall based on a replay of the eye-movements and the revision process (Van Gog et al., 2005).

Another possible avenue for future research is to focus on a single text section to understand and explore the relationship between multiple feedback and text revision variables in more depth. For instance, to get a more complete picture of how and why students use feedback, we propose to relate students' evaluations of the feedback to their feedback literacy and attitudes as well as to the characteristics and context of the feedback message (i.e., content, mode etc.). This multidimensional information is crucial to further improve theories on effective strategies for feedback use (cf. Carless & Boud, 2018; De Kleijn, 2021).

5. Implications and conclusion

The two present studies demonstrate individual differences in feedback processing strategies. Gaining insight in these different strategies can help teachers and students to get a better understanding of why and how feedback is used or not. This knowledge can be used to develop more effective feedback interventions that aim to increase the depth with which students read, evaluate, and use feedback for revision. One potentially powerful way to teach students how to engage with feedback more effectively might be to show and discuss examples of how other students and/or expert writers process written feedback and revise accordingly, using the replay of eye movement data in combination with thinking-aloud protocols and text analysis. This provides students with modeling examples of different feedback engagement strategies that they can imitate in their own writing and rewriting. The effectiveness of eye movement modeling examples for learning has already been demonstrated in different educational domains (Jarodzka et al., 2017).

The results of the present research might also have implications for the way teachers provide feedback, as it has been shown that the focus of their feedback has an effect on what and how students revise. More particularly, it is recommended to prioritize higher order feedback (e.g., structure, content) over lower order feedback (e.g., spelling, grammar, conventions), as higher order feedback supports students in making more meaningful revisions, and hence to become better writers.

In conclusion, this research is one of the first that has provided empirical evidence for the existence of different feedback processing strategies in students. We have shown that a triangulation of online and offline measures can reveal the underlying (meta)cognitive processes of reading and using feedback for revision. This information importantly refines current theories on how, when, and why feedback works and can be used to design more effective feedback interventions. Therefore, we encourage researchers to continue using these different process measures in their research on effective feedback use.

CRediT author statement

Renske Bouwer: Conceptualization, Methodology, Formal analysis, Investigation, Methodology, Writing – Original Draft, Review & Editing, Visualisation, Supervision.

Kim Dirkx: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Data gathering, Writing – Original Draft, Review & Editing, Project administration, Supervision.

References

Alqassab, M., Strijbos, J.-W., & Ufer, S. (2018). The impact of peer solution on peer-feedback provision on geometry proofs: Evidence from eye-movement analyses. *Learning and Instruction, 58*, 182–192. https://doi.org/10.1016/j.learninstruc.2018.07.003

Beason, L. (1993). Feedback and revision in writing across the curriculum classes. *Research in the Teaching of English, 27*, 395–421. https://doi.org/10.1007/s10212-017-0343-z

Berndt, M., Strijbos, J. W., & Fisher, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education, 33*, 31–49.

Bolzer, M., Strijbos, J. W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eyetracking: Role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning, 31*, 422–434.

Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (2018). *Developing evaluative judgment in higher education: Assessment for knowing and producing quality work*. Routledge.

Boud, D., & Molloy, E. (2013). *Feedback in higher and professional education*. Routledge.

Burke, D. (2009). Strategies for using feedback students bring to higher education. *Assessment & Evaluation in Higher Education, 34*(1), 41–50 (å).

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Cohen, A. D. (1987). Student processing of feedback on their compositions. In A. Wenden, & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 57–69). Prentice-Hall.

Cohen, A. D., & Cavalcanti, M. C. (1990). Feedback on compositions: Teacher and student verbal reports. In B. Kroll (Ed.), *Second language writing: Research insights from the classroom* (pp. 155–177). Cambridge University Press.

Dalmeijer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics, 23*(205). https://doi.org/10.1186/s12859-022-04675-1

De Kleijn, R. A. M. (2021). *Supporting student and teacher feedback literacy: An instructional model for student feedback processes. Assessment & Evaluation in Higher Education*. Advance online publication. https://doi.org/10.1080/02602938.2021.1967283

Elving, K. (2019). *Effectieve leeractiviteiten voor het schrijfonderwijs in havo 4 [Effective learning activities in secondary education writing classes]. Unpublished doctoral dissertation*. Amsterdam: LOT publishers.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analyses: Verbal reports as data*. The MIT Press.

Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition & Communication, 32*, 400–414. https://doi.org/10.2307/356602

Ferris, D. R. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *Tesol Quarterly, 29*, 33–53. https://doi.org/10.2307/3587804

Fitzgerald, J. (1987). Research on revision in writing. *Review of Educational Research, 57*(4), 481–506. https://doi.org/10.3102/00346543057004481

Graham, S. (2018). Instructional feedback in writing. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 145–168). Cambridge University Press.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Vol. 2. Reading, writing, and language processing* (pp. 176–240). Cambridge University Press.

Holmqvist, K., Nyström, N., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (Eds.). (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2018). *Multilevel analysis: Techniques and applications* (3rd ed.). New York: Routledge.

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing, 7*(3), 255–286. https://doi.org/10.1016/S1060-3743(98)90017-0

Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä (Ed.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Elsevier.

Jarodzka, H., Gruber, H., & Holmqvist, K. (2017). Eye tracking in Educational Science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research, 10*(1), 1–18. https://doi.org/10.16910/jemr.10.1.3

Jonsson, A., & Panadero, E. (2018). Facilitating students' active engagement with feedback. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 531–553). Cambridge University Press.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Khodabaks, F. S. (2020). *Reviseren naar aanleiding van geschreven feedback [Revising in response to written feedback*. Amsterdam: Vrije Universiteit [Unpublished Bachelor's thesis].

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication, 30*(3), 358–392. https://doi.org/10.1177/0741088313491692

McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology, 67*(3), 267–277. https://doi.org/10.1016/j.jclinepi.2013.08.015

McCurdy, P. (1992). What students do with composition feedback. In *Paper presented at the 27th annual TESOL convention, vancouver, BC, Canada*.

Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education, 45*(4), 527–540. https://doi.org/10.1080/02602938.2019.1667955

Nicol, D. (2020). *The power of internal feedback: Exploiting natural comparison processes. Assessment & evaluation in higher education*. Advance online publication. https://doi.org/10.1080/02602938.2020.1823314

Paulson, E. J., Alexander, J., & Armstrong, S. (2007). Peer review reviewed: Investigating the juxtaposition of composition students' eye movements and peer-review processes. *Research in the Teaching of English, 41*, 304–335.

Paulus, T. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing, 8*(3), 265–289. https://doi.org/10.1016/S1060-3743(99)80117-9

Pokorny, H., & Pickford, P. (2010). Complexity, cues and relationships: Student perceptions of feedback. *Active Learning in Higher Education, 11*(1), 21–30.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Robinson, M. D., & Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin, 27*(11), 1520–1532.

Sommer, R. (1968). The Hawthorne dogma. *Psychological Bulletin, 70*(6, Pt.1), 592–595. https://doi.org/10.1037/h0026728

Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. *College Composition & Communication, 31*(4), 378–388.

Sutton, P. (2012). Conceptualizing feedback literacy: Knowing, beginning, and acting. *Innovations in Education & Teaching International, 49*(1), 31–40. https://doi.org/10.1080/14703297.2012.647781

Treglia, M. O. (2009). Teacher-written commentary in college writing composition: How does it impact student revisions. *Composition Studies, 37*(1), 67–86.

Van Gog, T., Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior, 25*, 325–331. https://doi.org/10.1016/j.chb.2008.12.021

Van Gog, T., Paas, F., van Merriënboer, J., & Witte, P. (2005). Uncovering the problem-solving process: Cued-retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

Verhavert, S., Bouwer, R., Donche, V., & Maeyer, S. D. (2019). A meta-analysis on the reliability of comparative judgement. *Assessment in Education, 26*(5), 541–562. https://doi.org/10.1080/0969594x.2019.1602027.

Wengelin, A., Frid, J., & Johansson, R., & Johansson, V. (2019). Combining keystroke logging with other methods: Towards an experimental environment for writing process research. In E. Lindgren, & K. P. H. Sullivan (Eds.), *Observing writing* (pp. 30–49). Brill.

Williams, J., & Kane, D. (2009). Assessment and feedback: Institutional experiences of student feedback, 1996-2007. *Higher Education Quarterly, 63*(3), 264–286. https://doi.org/10.1111/j.1468- 2273.2009.00430.x

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*(1), 17–37. https://doi.org/10.1080/00461520.2016.1207538

Winstone, N. E., Nash, R. A., Rowntree, J. R., & Parker, M. (2017). It'd be useful, but I wouldn't use it: Barriers to university students' feedback-seeking and recipience. *Studies in Higher Education, 42*, 2026–2041. https://doi.org/10.1080/03075079.2015.1130032

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*(3087), 1–14. https://doi.org/10.3389/fpsyg.2019.03087