

University of Groningen

Rectifying Errors

Aben, Jochem

DOI:
[10.33612/diss.208729687](https://doi.org/10.33612/diss.208729687)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Aben, J. (2022). *Rectifying Errors: a Reconceptualization of the Role of Errors in Peer-Feedback Provision and Processing*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.
<https://doi.org/10.33612/diss.208729687>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

5

Dealing with Errors while Providing and Processing Peer-Feedback on Writing: A Mixed-Methods Approach

Aben, J.E.J., Mascareño Lara, M., Timmermans, A.C., Dingyloudi, F., & Srijbos J.W. (2022). *Dealing with errors while providing and processing peer feedback on writing: A mixed-methods approach*. Manuscript submitted for publication.

Abstract

Although previous research has stressed the central role of dealing with errors during feedback processes, this appears contingent on the interpersonal relationship between the feedback provider and feedback recipient due to the subjective nature of what constitutes an error. Therefore, this study investigates (1) which cognitive sub-phases can be distinguished during the process of dealing with errors while providing peer-feedback and processing peer-feedback, and (2) to what extent dealing with errors while providing peer-feedback and processing peer-feedback is affected by interpersonal factors.

Dyads of Dutch 11th grade students ($N = 12$, age range = 16-18) provided and processed peer-feedback on an argumentative text while thinking-aloud, and they reflected on the process during an interview afterwards. The think-aloud utterances and interviews were transcribed, and analyzed with a mixed-methods, exploratory sequential design, applying a quantitative content analysis, and a qualitative thematic analysis.

The process of dealing with errors during peer-feedback provision displayed two patterns. In the first pattern, error identification seemed to happen simultaneously with the decoding, and often any evaluation-related thoughts were lacking. During the second pattern, error-identification seemed to occur as a result of an interpreting/evaluating phase. Also during peer-feedback processing, two main patterns were visible: students either knew immediately whether they agreed with received feedback, or they first had to study the feedback more thoroughly. Additionally, the interpersonal factors 'perceived language skills' and 'perceived effort' seemed to affect students only implicitly during feedback provision, whereas interpersonal factors seemed to affect the majority of the students explicitly during feedback processing.

Keywords

peer-feedback; interpersonal factors; argumentative essay writing; think-aloud

Previous research in education has stressed the central role of dealing with errors during feedback processes (e.g., Aben et al., 2019; Fong et al., 2018; Timms et al., 2016). Errors, often defined as deviations from a norm (e.g., Gloy, 1987; Oser & Spychiger, 2005; Rach et al., 2012; Spychiger et al., 2006), can be considered as fundamental prerequisites for learning. After all, students that process feedback are likely to be confronted with (a) performance elements perceived as erroneous or improvable by the feedback provider, and (b) feedback elements identifying and criticizing these performance elements. Since error-making and problem-solving are crucial for learning, and feedback is likely to function as a scaffold to reduce the gap between the current and a desired performance (Ramaprasad, 1983), errors are viewed as opportunities for learning and play a central role in the provision and processing of feedback (Fong et al., 2018).

A consequence of viewing errors as deviations from a norm, is that the notion of what constitutes an 'error' becomes intersubjective, because the notion of a 'norm' is often ambiguous. Norms may refer to (a) implicit, personal codes of conduct, e.g., situation-specific behavioral expectations (Lapinski & Rimal, 2005); (b) generalized written rules, e.g., grammar prescriptions in a specific language (Zwicky, 1980); (c) the minimum performance for a passing grade, e.g., examinations (e.g., Prince, 2016); or (d) anything in between. Consequently, people may have different opinions regarding the extent to which performances, or parts thereof, deviate from norms and can therefore be considered erroneous or improvable (Zwicky, 1980).

The subjective character of errors also implies that feedback processes are contingent on the interpersonal relationship between the feedback provider and feedback recipient (Aben et al., 2019; Esterhazy & Damşa, 2019). That is, the feedback provider's perception of the quality of the recipients' performance, or the feedback recipient's perception of the adequacy of the provided feedback, may depend on various interpersonal factors such as feedback providers' and recipients' experiences with each other, and perceptions of each other (Strijbos & Müller, 2014; Winstone et al., 2017). In particular when students provide and receive peer-feedback, interpersonal factors appear to affect the feedback processing (e.g., Alqassab et al., 2018b; Berndt et al., 2018; Strijbos et al., 2010); presumably because students may be more likely to doubt the expertise of a peer (horizontal constellation), than, for example, a teacher (vertical constellation) (Strijbos & Müller, 2014). These effects are particularly relevant when predicting peer-

feedback reception and explaining the outcomes of peer-feedback processing (Aben et al., 2019; Esterhazy & Damşa, 2019).

Nevertheless, empirical research investigating what the process of dealing with errors looks like and how it may be affected by interpersonal factors during the provision and processing of feedback, is scarce (Máñez et al., 2019). A better understanding of such feedback processes is necessary to improve our comprehension of how feedback processing and outcomes can be explained and eventually improved (Handley et al., 2011). Therefore, we study the processes of peer-feedback provision and feedback processing in the domain of argumentative essay writing, a domain where peer-feedback is a frequently used method to enhance students' writing skills and performance (Double et al., 2020; Hoogeveen & Van Gelderen, 2013; Huisman et al., 2019).

The Role of Errors in Feedback Provision and Processing

The role of errors takes a central place in a handful conceptualizations of (peer-)feedback provision and/or processing (e.g., Aben et al., 2019; Fong et al., 2018; Timms et al., 2016). For instance, Aben et al. (2019) assume the existence of cognitive sub-phases in the process of dealing with errors while providing and processing feedback. Regarding the feedback provision phase, Aben et al. (2019) take a performance that potentially contains errors as starting point. For example, in the context of (peer-)feedback on a written text, feedback providers may identify an error, which refers to the moment an individual observes that a text element does not meet a norm. Subsequently, feedback providers may decode the identified error, i.e., assign meaning to the error, by labeling the error as an error of a particular kind (e.g., grammar error, spelling error, argumentation error). Third, this decoding may lead to the evaluation of an error, which refers to the assessment of what characteristics make a text element erroneous and/or to the thinking about how an observed error may be improved. Error evaluation may be followed by the encoding of a feedback remark on the specific error, i.e., the translation of the interpretation and evaluation of the error into the production of verbal and/or nonverbal signs. Finally, this remark may be sent to the feedback recipient in the form of a feedback remark.

Feedback recipients, on the other hand, first have to read a provided feedback remark in relation to their written text. Thereafter, they have to decode the feedback remark, i.e., assign meaning to it in order to interpret the feedback (Akin et al., 1970). Similar to the feedback provision phase, the decoding of the feedback remark may

lead to the evaluation of the feedback remark, which refers to the activity of deciding to what extent one agrees with the feedback and/or the potential acknowledgement that one made an error. Finally, an output (e.g., correction of an error of a particular kind, such as a grammar error, spelling error, or argumentation error) may be encoded, potentially (partly) based upon the feedback remark. This output may express disagreement with the feedback or may show the intention to act upon the feedback and to correct the error. If the feedback recipient acts upon feedback, this leads to a revised performance, which ideally implies that the initial erroneous performance is (partially) rectified (Aben et al., 2019).

The existence of cognitive sub-phases in the process of dealing with errors while processing feedback was also found by Ahmadian et al. (2019), who collected think-aloud utterances related to dealing with errors of university students while they processed teacher feedback on writing performance, in the context of second language acquisition. The analyses indicated that their students displayed a recurrent pattern of dealing with grammatical errors. This pattern existed of the subsequent phases of reading the sentence in their own text (e.g., "...type of diversity which reflect..."), reading the feedback ("third person"), encoding a revised output ("which reflects"), and explaining the grammatical error ("I should have used 's' here for the third person"). This pattern differs from the sequence of sub-phases as hypothesized by Aben et al. (2019), in the sense that evidence of a feedback decoding phase was not found by Ahmadian et al. (2019), and that in their results the evaluation sub-phase takes the form of a retrospective error explanation.

These different configurations of the process of dealing with errors while providing and processing feedback may respond to the type(s) of errors that must be dealt with (Ahmadian et al., 2019; Kim & Bowles, 2019). Kim and Bowles (2019) performed a think-aloud study in the context of second language acquisition, that aimed to capture feedback processing. They investigated, among other matters, whether the type of error identified by the feedback provider related to feedback recipients' depth of feedback processing. The results indicated that students took more time to process, interpret, and evaluate the feedback on errors related to higher order concerns (e.g., content, argumentation, and paragraph structure) than on errors related to lower order concerns (e.g., spelling, grammar, and interpunctuation).

The central role that errors play in feedback provision and feedback processing may also be reflected in the way learners deal

with pluses. In the current study, pluses are defined as performance elements perceived by either the feedback provider or feedback recipient as meeting or surpassing a norm. For example, Máñez et al. (2017) compared the time that students took to process automatically generated feedback on errors with the time students took to process automatically generated feedback on pluses. During a reading task, students had to answer multiple-choice questions and select the relevant textual information on which they based their answer. They found that students spend more time to process the feedback that was provided when they did not execute the task successfully (i.e., errors) than to process the feedback that was provided when they did execute the task successfully (i.e., pluses). According to Máñez et al. (2017), these results may suggest that students are (intuitively) aware of the importance of understanding errors in feedback processing in order to enhance their skills and performance.

Interpersonal Factors and Peer-Feedback Processes

In addition to the assumption that the process of dealing with errors in peer-feedback may contain different sub-phases, the process of dealing with errors may be affected by interpersonal factors, i.e. factors describing the relationship between people (Aben et al., 2019; Esterhazy & Damşa, 2019; Winstone, 2017). These interpersonal factors are gradually shaped over time through past experiences of the actor with the same partner and consequently the actor's developing perception of the partner (Gibson, 1969; Upshaw, 1978). For example, students that process peer-feedback are members of social constellations, such as classrooms, for extended periods of time. Hence, they are likely to attend group discussions, to collaborate with the same peers on learning tasks, and they could be aware of their peers' skills or grades. Via their joint engagement in classroom activities, students collect pieces of information about their peers, which may either consciously or unconsciously, contribute to the composition of a mental representation of their peers' expertise—as well as their own expertise in comparison to their peers' expertise— and may, thus, potentially affect feedback processing (Aben et al., 2019; Strijbos & Müller, 2014; Winstone et al., 2017).

Although, to our knowledge, as of yet no studies have investigated the effects of interpersonal factors on the process of dealing with errors in particular, at least three interpersonal factors received considerable attention regarding their role in peer-feedback provision

and/or processing in general. First, several studies found that friendship could lead to biases in peer-feedback or peer-grading (e.g., Harris & Brown, 2013; Cheng & Warren, 1997; Panadero et al., 2013). Students may invest more effort in providing peer-feedback when they consider the peer a friend than when they do not consider the peer a friend (Finkelstein, Fishback, & Tu, 2017).

Second, the effort peers devote to providing and/or processing peer-feedback may play a role in peer-feedback processing. For example, Timmers et al. (2020) found that the effort invested in feedback processing was predicted by an individual's task-value beliefs, such as the importance of and interest in the task. As these task-value beliefs are situation and consequently task-dependent (Eccles & Wigfield, 2002), this may imply that the perceived effort that a peer has devoted to providing feedback, as perceived by the feedback recipient, may also affect the recipient's peer-feedback processing.

Third, students' peer-feedback processing may be affected by perceptions of their peers' language skills (e.g., Berndt et al., 2018; Strijbos et al., 2010). For example, Berndt et al. (2018) manipulated perceived language skills by giving their students scenarios containing essays written by a fictional student who was provided with fictional peer-feedback. They found that feedback provided by a peer with low competence was perceived as less adequate than feedback provided by a highly competent peer. Also in Chapter 4 of this dissertation the perceived language skills of peer-feedback providers were manipulated. The results showed that the perceived language skills of the feedback provider significantly related to the proportion of textual revisions made by students based on feedback related to writing style: students revised their text more often in line with this type of feedback when they thought it was provided by a peer perceived to have stronger language skills than their own than when they thought it was provided by a peer perceived to have weaker language skills than their own.

The Current Study

Empirical research investigating what the process of dealing with errors looks like, and how interpersonal factors affect the process of dealing with errors while providing and processing peer-feedback is currently scarce. This study adds to the literature by investigating (1) which cognitive sub-phases can be distinguished during the process of dealing with errors while providing peer-feedback (i.e., identifying, decoding, evaluating, and encoding) and processing peer-feedback (i.e., reading,

decoding, evaluating, and revising), and (2) to what extent dealing with errors while providing and processing peer-feedback is affected by interpersonal factors (e.g., friendship, perceived skills, perceived effort). We adopted a mixed-methods design, triangulating data from a quantitative content analysis as well as a qualitative thematic analysis of think-aloud data utterances and semi-structured interviews (Dingyloudi & Strijbos, 2018b).

Method

Design

We conducted a study in which data of think-aloud protocols and semi-structured interviews were triangulated. Dyads of students provided and processed peer-feedback on argumentative texts while thinking-aloud, and reflected on the processes during an interview afterwards. After the data collection, the think-aloud utterances and interviews were transcribed by two research-assistants, and analyzed with a mixed-methods exploratory sequential design, first applying a quantitative content analysis (top-down) and second a qualitative thematic analysis (bottom-up) (Dingyloudi & Strijbos, 2018b). We integrated the top-down and bottom-up analyses to distinguish the sub-phases in the process of dealing with errors in the context of peer-feedback provision and processing, and the role of interpersonal factors in this process.

Context and Participants

Data were collected at a public high school in the Northern region of the Netherlands in the context of the subject 'Dutch language and literature', which focuses on writing, reading, presenting and argumentation skills, among others. During a period that was devoted to writing skills, the teachers taught the students about the structure, genre characteristics and quality criteria of an argumentative text. Additionally, students produced an argumentative text, received peer-feedback, received teacher feedback, and rewrote their first drafts based on the feedback. Students from all four 11th grade classes at this school ($N \approx 100$, age range = 16-18) were asked by their four teachers whether they wanted to participate in a study about peer-feedback. Participation in this study was voluntary. Twelve students from two of the 11th grade classes (class A: $n = 7$; class B: $n = 5$) agreed to participate. Among those students

were nine boys and three girls (Age: $M = 16.6$ years; $SD = 0.95$; range = 16–18 years). They all voluntarily signed a form of active informed consent.

In order to increase the ecological validity, the data collection blended in with the high school's planning: Students provided and processed feedback on the argumentative texts they had to write as part of their regular school activities. Additionally, the data collection took place in the same week in which the non-participating 11th grade students had to hand in their written texts. The method was approved by the ethical committee of the University of Groningen prior to data collection.

Dyad Composition, Procedure and Instruments

Data collection took place on two research days. On the first research day, students produced an argumentative text in which they had to defend the point of view that self-driving cars would make the traffic safer in the future. This text had to be about 400 words long and had to be divided into an introduction, body, and conclusion section. Additionally, they filled out a questionnaire about the interpersonal factor 'perceived language skills'. Perceived language skills were measured by asking: "How good are your classmates in the school subject Dutch, compared to you?" The sentence was followed by a list of all participating classmates for whom the students had to indicate whether they perceived the language skills of each classmate as *a lot worse than mine* (1), *a bit worse than mine* (2), *about as good as mine* (3), *a bit better than mine* (4), *a lot better than mine* (5), or *I don't know who this student is*.

Subsequently, six dyads were composed, based on two criteria. The first criterion was that the students in a dyad knew each other, so that an interpersonal relationship could be assumed. Second, in order to optimize variability across dyads, the dyads were composed to represent different combinations of perceived Dutch language skills between their members. Table 5.1 shows a descriptive overview of the dyads (all names are pseudonymized).

On the second research day, two students simultaneously met one of the researchers at school, who brought them to two separate rooms in the school, each together with one of the two research-assistants. The students simultaneously provided feedback on each other's draft texts. Subsequently, the provided feedback was exchanged and the students processed the feedback that had just been provided. These

sessions were audio-recorded, lasted between 80 and 90 minutes per dyad, and took place between four and six days after the students had to hand in their draft text for the writing assignment as part of their curriculum for the subject ‘Dutch language and literature’.

Table 5.1. Description of dyads based on the two criteria for dyad composition.

Dyad number	Name A	Name B	Criterion 1. Do they know each other?	Criterion 2. Perceived language skills ¹
1	Mary	James	Yes	Strong – Strong
2	Robert	John	Yes	Weak – Weak
3	Linda	Michael	Yes	Strong – Weak
4	Steven	Sarah	Yes	Weak – Strong
5	David	Thomas	Yes	Strong – Strong
6	Daniel	Anthony	Yes	Weak – Weak

¹ A perceives B as... - B perceives A as...

The research-assistant welcomed the student to the room and instructed students to “provide feedback on [their peer]’s text, to the best of [their] ability, with the aim to help the writer to improve the text”. The research-assistant provided and explained a brief overview with potential quality criteria to provide feedback on (i.e., spelling, writing style, argumentation, text structure; see Supplemental materials, Chapter 5, section 1). Data were collected by means of a concurrent non-metacognitive think-aloud procedure while students provided and processed peer-feedback, implying that students had to think aloud while providing and processing feedback, but did not have to reflect on their thinking. This type of think-aloud procedure was chosen, as it decreases the chance that participants feel the urge to argue their thinking and hence interrupt their thought process, compared to a concurrent meta-cognitive think-aloud procedure (Ericsson & Simon, 1998).

Students were acquainted with thinking aloud by providing feedback on a practice text (Bowles, 2010), using *Microsoft Word’s Track Changes* and *Comment* functions. Hereafter, students had 25 minutes to provide feedback on their peer’s text ($M_{text\ length} = 492$ words, $SD_{text\ length} = 171$ words), while explicating their thoughts. Every time they remained quiet for five seconds, the research-assistant asked “what are you thinking at this moment?” (Bowles, 2010; Máñez et al., 2019). After 25 minutes, the research-assistant saved the texts with feedback on USB-drives and exchanged them. Then, the students had 25 minutes to process the feedback that had been provided by the peer. The students were told that they had the freedom to ignore feedback if they wanted to, as they remained the owners of their text. They were instructed to revise their text as if they had to hand in the final version at the end of the session.

The session ended with the semi-structured interview, conducted by the research-assistant. First, students were asked to reflect on the way they dealt with errors during all sub-phases of peer-feedback provision and processing. Second, students were asked to reflect on the way they had been affected by the interpersonal relationship with the peer. The research-assistant initially asked non-guided open questions providing the students the opportunity to mention potential effects of the interpersonal relationship themselves, before asking more guided questions (see Supplemental materials, Chapter 5, section 2).

Data Analysis

Quantitative Content Analysis. Content analysis enables replicable and valid inferences from data to their context (Krippendorff, 2004), and converting qualitative statements into quantitative data (Stemler, 2015). We applied multi-valued coding using Atlas.ti (2018) version 8, implying that a quotation could receive a code from different semantic domains. We distinguished six semantic domains: (1) sub-phases in dealing with errors while providing feedback (identifying, decoding, evaluating, encoding); (2) sub-phases in dealing with pluses while providing feedback (identifying, decoding, evaluating, encoding); (3) sub-phases in dealing with errors while processing feedback (reading, evaluating, revising); (4) sub-phases in dealing with pluses while processing feedback (reading, evaluating, revising); (5) mentioning the dyad partner (‘he/she’, ‘[name dyad partner]’, ‘you’); and (6) effects of interpersonal relationship (friendship, perceived language skills, perceived skills general, perceived effort, no effort). While coding the

think-aloud utterances (i.e., utterances during the feedback provision and feedback processing phases) the leading question for the coders was: what is the student *doing*? While coding the interview transcripts, the leading question for the coders was: what phase is the student *talking about*?

The coding scheme was developed by the first author, and revised after initial trials performed by the first author and the two research-assistants. We made two revisions that require further elaboration. First, the error decoding and plus decoding sub-phases could not be coded in the feedback processing phase, because it was problematic to distinguish the reading and decoding phases based upon think-aloud utterances. That is, we could only observe evidence of students reading the feedback, making a valid empirical distinction between the two processes of reading and decoding impossible.

Second, we initially had not conceived semantic domain (5) as part of the coding scheme; however, we added this domain because it can be argued that linguistic references to the dyad partner form evidence of feedback providers’ awareness of the feedback recipient, and hence of the interpersonal relationship between the feedback recipient and the self. Semantic domain (5) was only coded for the think-aloud data; all other semantic domains were coded for both the think-aloud data and the interview data (see Table 5.2).

Power Analysis. A power analysis revealed that the number of twelve students was likely to be sufficient to calculate a reliable interrater reliability (Krippendorff’s α) for semantic domain (1) ‘dealing with errors’. Semantic domain (1) contained four categories that were expected to be equally distributed. That is, it was expected that, if students identified an error, they subsequently decoded, and evaluated the error, and encoded a feedback remark. Applying equation 1 showed that a code should occur at least 122 times within the semantic domain ‘dealing with errors’ (Atlas.ti, 2018), with N_c = Number of coders = 3; $Z_p^2 = z^2$ belonging to the highest acceptable p -value (.05) = $1.65^2 = 2.72$; and P_c = chance for one category within one semantic domain = $1/\text{number of categories within domain} = 1/4 = .25$.

$$(N_c * Z_p^2) * (((1 + \alpha_{min}) * (3 - \alpha_{min})) / ((4 * (1 - \alpha_{min})) * (P_c * (1 - P_c)))) - \alpha_{min}) \quad (1)$$

Table 5.2. Description of the semantic domains and codes used in the quantitative content analysis.

Semantic domain	Code	Description	Example
(1) Sub-phases in dealing with errors while providing feedback	Identifying	Observing that a text element does not meet an expectation.	“But there is a period, and I don’t think it should.”
	Decoding	Assigning meaning to an error, by explicitly categorizing a text element as error of category X.	“Most are spelling mistakes, or typos.”
	Evaluating	Assessing what characteristics make a text element an error or elaborating on how an observed error may be improved.	“...because yes, it is simply said that it was not seen as literature, but why not is not clear to me.”
	Encoding	Creating (thinking about, formulating and/or typing) feedback as a result of a text element that is viewed as negative.	“I would make this two sentences.”
(2) Sub-phases in dealing with pluses while providing feedback	Identifying	Observing that a text element meets an expectation.	“I think it’s quite a good introduction.” “I think it is, um, a good point of view.”
	Decoding	Assigning meaning to a plus, by explicitly categorizing a text element as plus of category X.	“His writing style is good.” “Lay-out is completely correct.”
	Evaluating	Assessing what characteristics make a text element a plus.	“The, um, topic is introduced and he takes a clear position, that is indeed what is needed.” “Clear statement, main arguments with examples.”
	Encoding	Creating (thinking about, formulating and/or typing) feedback as a result of a text element that is viewed as positive.	[writes down:] “make this into two separate sentences.”

Table 5.2. (Continued).

(3) Sub-phases in dealing with errors while processing feedback	Reading	Assigning meaning to feedback that was intended by the feedback provider as feedback on an error, in order to interpret the feedback	Reading the feedback that is written by the feedback provider
	Evaluating	Stating or determining to what extent one agrees with the feedback that was intended by the feedback provider as feedback on an error.	<p>“Yes, I do agree.”</p> <p>“Yes, I would do that too.”</p> <p>“But I don’t really agree with that, because I think this is just a good sentence.”</p>
	Revising	Creating (thinking about, formulating and/or typing) a textual alternative as a result of a text element that feedback that was intended by the feedback provider as feedback on an error.	“I’m trying to come up with a new title, kind of a creative one.”
(4) Sub-phases in dealing with pluses while processing feedback	Reading	Assigning meaning to feedback that was intended by the feedback provider as feedback on a plus, in order to interpret the feedback	Reading the feedback that is written by the feedback provider
	Evaluating	Stating or determining to what extent one agrees with the feedback that was intended by the feedback provider as feedback on a plus.	“of course he also uses a quote there, and um, something that is also known to me, so that draws the attention.”
	Revising	Creating (thinking about, formulating and/or typing) a textual alternative as a result of feedback that was intended by the feedback provider as feedback on a plus.	“...so with that comment, I don’t really need to do anything. Then I can just delete that comment, right?”

Table 5.2. (Continued).

(5) Mentioning the dyad partner	He/she	Reference to the other with ‘he’ or ‘she’.	“ <u>H</u> e is right about this.”
	Name	Explicit mentioning of the other’s name.	“I thought that was a good sentence to grab the attention; <u>Robert</u> thought the same.”
	You	Reference to the other with ‘you’.	“This is not possible. I guess <u>you</u> mean X.”
(6) Effects of interpersonal relationship	Friendship	Remarks about friendship or the extent to which the one knows the other.	“I don’t really know Thomas that well.”
	Perceived language skills	Remarks about perceived language skills.	“Because she is good at Dutch.”
	Perceived skills general	Remarks about perceived skills in general.	“She is just generally smart.”
	Perceived effort	Remarks about perceived effort put in task.	“She puts little energy and time into it.”
No effect		Explicitly stating that the interpersonal relationship did not affect feedback provision or processing behavior.	“I just looked at the feedback and I don’t really care what person wrote it.”

Hence, the least number of errors to be identified by students while providing feedback for an interpretable interrater reliability was 122 (codes within domain) / 4 (categories within domain) / 12 (students) = 2.5 errors. This was deemed likely.

Hereafter, a subsample of 16.7% of the complete data set was selected and coded by the first author and the two research-assistants in order to calculate an interrater reliability. This subsample of 16.7% contained two feedback provision think-aloud transcripts, two feedback processing think-aloud transcripts, and two semi-structured interview transcripts, of six individual students, that were not used in the practice rounds.

The interrater reliability between the three coders based on the 16.7% of the data was sufficient for all six semantic domains (Krippendorff's α range: $.70 - .81$). This was also the case when the interrater reliability was calculated between only the two research-assistants, leaving out the first author's codings (Krippendorff's α range: $.73 - .83$). Subsequently, the two research-assistants coded the remainder of the data.

Qualitative Thematic Analysis. Hereafter, a thematic analysis of the think-aloud data and interview data was performed by the first author and one research-assistant. Thematic analysis is a method to identify, analyse, and report patterns within data (Braun & Clarke, 2006). The aim of the qualitative analysis was to make an elaborate interpretation of the way students dealt with errors and the way they were affected by interpersonal factors, allowing us to move beyond the predefined codes of the quantitative content analysis. Instead of counting the occurrence of codes we expected to encounter (i.e., top-down), we extracted patterns from the data (i.e., bottom-up).

The first author and the research-assistant re-listened the audio-recordings and noted down their thoughts related to dealing with errors and effects of interpersonal factors, also when these thoughts were based upon students' utterances before, during, or after the feedback provision, feedback processing and interview activities. They did this, independently, for two of the twelve students. The first author produced a list of themes, which was extended after a discussion with the research-assistant. Hereafter, the first author and the research-assistant independently re-listened the remaining audio-recordings and noted down their thoughts. In the end, they both categorized their own thoughts using the list of themes. The thematic analysis ended with a two-hour discussion between the first author and the

research-assistant, identifying and describing the main patterns derived from thematic analysis related to dealing with errors and effects of interpersonal factors.

Triangulation. Both types of data (i.e., think-aloud utterances and interviews) were necessary to address the research questions. Whereas the quantitative content analysis of the think-aloud utterances provided insights into the extent to which sub-phases of the process of dealing with errors occurred, the qualitative thematic analysis provided descriptions of what the sub-phases looked like. Similarly, the interviews provided enriching information about potential effects of interpersonal factors on the process of dealing with errors that may have been only partly observable in students' think-aloud utterances.

Results

The main patterns identified are categorized into four sections: (1) Dealing with errors while providing feedback; (2) Effects of interpersonal factors while providing feedback; (3) Dealing with errors while processing feedback; (4) Effects of interpersonal factors while processing feedback. Section (1) and (3) are only based upon the quantitative content analysis and the qualitative thematic analysis of the think-aloud data, as, retrospectively, these analyses provided more reliable and objective insights into the process of dealing with errors than the interview data. That is, in the interview students were not able to properly describe the way they had dealt with errors. Sections (2) and (4) are a synthesis of the quantitative content analysis and qualitative thematic analysis of both the think-aloud and interview data.

Dealing with Errors while Providing Feedback

The provision of feedback, as conducted by the students in the sample, could be characterized as linear. Students opened the text document, and immediately started reading the text from start to end. They wrote comments, related to particular text elements, at the moment they first encountered those text elements. When they had read the end of the text for the first time, some students started re-reading the text from beginning to end, others did the same only after being prompted by the research-assistant. During their 'second round', all students produced new feedback remarks, based on text elements other than those dealt

with during the ‘first round’.

During this procedure of providing feedback, the think-aloud utterances showed that students focused more on errors than on pluses. On the one hand, the pattern of identifying was mixed. Seven students more often identified pluses, and five students more often identified errors. Additionally, all students hardly declared to decode errors or pluses. On the other hand, the majority of students said more often that they were evaluating errors than pluses, and nearly all said to encode feedback based on errors more often than feedback based on pluses (Table 5.3). The emphasis on encoding feedback based on errors also became clear from the fact that several students asked whether they also ‘had to’ provide feedback on pluses. For example, Michael asked: “I also have to provide compliments, right?” (Michael, think-aloud).

In the identification of errors, students’ behavior consistently displayed one of two patterns that can be described as ‘bumping into errors’ and ‘discovering of errors’. The ‘bumping into errors’ pattern typically occurred in relation to lower order concerns, and was characterized by a phase of reading the text, immediately followed by the identification of an error, or even an immediate error correction. As such, the identification phase, potential decoding phase, and potential evaluation phase either did not occur, or occurred so quickly that students did not distinguish between them in their think-aloud utterances. The following example illustrates the pattern of bumping into errors: “[reads the text:] ‘this is a joke about not being able to speak English very well and that you will learn it.’ At least a comma has been forgotten before ‘and’. And that you will – oh I need to write that down” (David, think-aloud). This example shows that the first time David read the sentence written by his dyad partner, he immediately identified the textual element as erroneous: while reading, he bumped into the error, and almost simultaneously provided a correction in the form of an added comma, which resulted in the encoding of a feedback remark.

The intuitive nature of the pattern of bumping into errors contrasted with the pattern of discovering of errors, which typically occurred in relation to higher order concerns, and was the result of an elaborated interpretation process. During this interpretation process, students read a part of a text, realized that they did not fully grasp it, reread the same part, potentially in multiple iterations, and inferred what the author tried to say. This interpretation process led to the identification and evaluation of an error. For example:

Table 5.3. Frequencies for codes related to semantic domains (1) Dealing with errors while providing feedback, and (2) Dealing with pluses while providing feedback.

	Dealing with errors						Dealing with pluses											
	Identifying		Decoding		Evaluating		Encoding		Identifying		Decoding		Evaluating		Encoding		Total	
	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%
Linda	15	28.9%	0	0.0%	10	19.2%	13	25%	9	17.3%	2	3.9%	3	5.8%	0	0.0%	52	
James	17	22.7%	0	0.0%	17	22.7%	19	25.3%	14	18.7%	0	0.0%	6	8.0%	2	2.7%	75	
Mary	5	7.3%	2	2.9%	11	15.9%	32	46.4%	9	13.0%	1	1.5%	5	7.3%	4	5.8%	69	
John	27	33.3%	1	1.2%	14	17.3%	28	34.6%	6	7.4%	0	0.0%	2	2.5%	3	3.7%	81	
Michael	11	20.8%	1	1.9%	10	18.9%	18	34.0%	9	17.0%	1	1.9%	1	1.9%	2	3.8%	53	
Daniel	19	22.1%	0	0.0%	25	29.1%	23	26.7%	13	15.1%	0	0.0%	3	3.5%	3	3.5%	86	
Steven	8	16.7%	0	0.0%	4	8.3%	5	10.4%	15	31.3%	2	4.2%	5	10.4%	9	18.8%	48	
Anthony	12	18.2%	0	0.0%	10	15.2%	14	21.2%	16	24.2%	0	0.0%	7	10.6%	7	10.6%	66	
David	14	18.9%	0	0.0%	24	32.4%	17	23.0%	7	9.5%	0	0.0%	5	6.7%	7	9.5%	74	
Sarah	8	17.4%	0	0.0%	5	10.9%	28	60.9%	3	6.5%	0	0.0%	0	0.0%	2	4.4%	46	
Thomas	11	15.3%	0	0.0%	9	12.5%	12	16.7%	24	33.3%	0	0.0%	11	15.3%	5	6.9%	72	
Robert	0	0.0%	0	0.0%	1	3.9%	16	61.5%	5	19.2%	0	0.0%	0	0.0%	4	15.4%	26	

[reads the text] This was the ehm, (4) counter-argument it looks like, but then, this is the conclusion. [...] Ooh. It is placed again over here. Never mind. I missed the final page, but it is alright like this [rereads same part of the text] Yes, she uses the argument that one [...] keeps learning Dutch [...] But then one could use the same argument in favor of that one learns English in this way [...] Ehm. I will write that down now (James, think-aloud).

In this example, the error, which could be characterized as an argumentation error, was identified after an extensive thinking process. First, James did not understand what he had read, then he reread the same part; he first believed the author was right, but only in the end he identified and evaluated an error.

These two patterns of identifying errors were similar in the sense that comparisons between the text and quality criteria often played a role. Eight of the twelve students (Mary, James, John, Michael, David, Thomas, Daniel, Anthony) repeatedly used their knowledge about what an argumentative text should look like (i.e., quality criteria) during their processes of identifying and evaluating errors, and encoding feedback remarks. For example:

[Teacher name; omitted] always comes up with three, ehm, rules that [the title of an argumentative text] should follow, being that a title should be covering, explaining, and catchy. Covering and explaining, that's kind of fine, but catchy... I would, ehm, a bit more, ehm... (Anthony, think-aloud).

This excerpt illustrates that the comparison with criteria contributed to the identification of errors, in this example the discovering of an error, by offering a ground for a proper evaluation.

Effects of Interpersonal Factors while Providing Feedback

There was a clear difference between the extent to which students *declared* in the interview to have been affected by interpersonal factors while providing feedback and the extent to which they *seemed* to have been affected during the think-aloud while providing feedback. On the one hand, ten of the twelve students (all, except John and Anthony) strongly argued in the interviews that they were not affected by the interpersonal relationship with the feedback recipient while providing

feedback. They shared the view that they 'just' provided feedback on a text, and that they would have done that the same way when they would have had to provide feedback on a text written by another peer: "It's a text, not a person what I read" (Robert, interview); "[The way in which I provide feedback] would not differ that much [from] if I had to provide feedback to someone else" (Thomas, interview).

On the other hand, half the students (Mary, James, Steven, Sarah, Daniel, Anthony) showed awareness of the feedback recipient, without being prompted by one of the research-assistants. This happened before, in between, or after the think-aloud activities, during the interviews, and also during the think-aloud activities. These students seemed to have taken into account the perceived language skills, and/or the perceived effort of the feedback recipient. A few examples: James mentioned, just before he started providing feedback, that "Mary [Mary = the feedback recipient] has in general, let's say, a higher level of writing skills, so [laughs] (perceived language skills)." Sarah stated, before she started providing feedback: "He [He = the feedback recipient] also just texted me and he said, like, 'I wrote it late in the evening', so, don't worry too much about it (perceived effort)." Daniel said during the interview: "I know that he [he = the feedback recipient] is dyslectic himself. So you notice that, it pops up" (perceived language skills)." And Mary said, while providing feedback, "James [James = the feedback recipient] told me that he wrote this essay just quickly within 23 minutes (perceived effort)." Additionally, eleven of the twelve students (all, except Robert) explicitly referred in their think-aloud utterances to the feedback recipient by saying 'he (his)' or 'she (her)', 'you', and/or by mentioning the recipient's name (see Table 5.4). Examples are: "I believe his point of view is not clear in the text." (Linda, think-aloud); "I will write this down anyway: "is 'learnability' a word? Maybe he can use another word" (Mary, think-aloud).

Dealing with Errors While Processing Feedback

The processing of feedback, as conducted by the students in the sample, was less linear than the provision of feedback. First, all students made interchangeably use of two strategies. When they revised their text by taking the feedback as their starting point, they processed each of the feedback remarks one by one, and decided for each of them whether they deemed any textual revision necessary. When they took their texts as their starting point, they read their text and switched to a feedback remark when they encountered one. Second, eight students (all except

Table 5.4. Frequencies for codes related to semantic domain (5) mentioning the other, during the provision of feedback.

	He/she	Name	You	Total
Linda	7	0	0	7
James	13	1	0	13
Mary	22	2	12	36
John	15	0	0	15
Michael	6	0	0	6
Daniel	12	0	1	13
Steven	11	0	0	11
Anthony	36	0	0	36
David	29	0	12	41
Sarah	2	0	0	2
Thomas	20	0	0	20
Robert	0	0	0	0
Average	14.4	0.3	2.1	16.7

for James, Anthony, Sarah, and Thomas), repeatedly did not deal with feedback immediately as they read it. When they encountered a feedback remark asking for textual adaptations which they agreed with, but simultaneously did not know how to revise their text accordingly, they decided to return to this remark later.

During the feedback processing phase students focused more on errors than on pluses. Table 5.5 shows that all students voiced more often think-aloud utterances related to any of the sub-phases of dealing with errors (i.e., decoding, evaluating, revising) than to any of the same sub-phases related to dealing with pluses; except for Sarah who made more utterances regarding reading, evaluating, and revising for pluses than for errors. This did not seem to be only a logical consequence of their dyad partner’s focus on errors during the feedback provision phase, but several students additionally seemed to experience feedback on pluses as a distraction: “I will delete, eh, the positive remarks [...]. Then I can focus on the worse parts. Alright” (Daniel, think-aloud). All students hardly thought about pluses, which consequently only seemed to stick to a limited extent: “[reads comment:] ‘Good title, maybe a bit more general.’ Well, I can immediately delete that one. Ehm. Like this.” (Robert, think-aloud)

Table 5.5. Frequencies for codes related to semantic domains (3) Dealing with errors while processing feedback; (4) Dealing with pluses while processing feedback.

	Dealing with errors						Dealing with pluses							
	Reading		Evaluating		Revising		Reading		Evaluating		Revising		Total	
	f	%	f	%	f	%	f	%	f	%	f	%	f	%
Linda	18	36.7%	14	28.6%	15	30.6%	2	4.1%	0	0.0%	0	0.0%	49	0.0%
James	23	30.3%	18	23.7%	29	38.2%	5	6.6%	1	1.3%	0	0.0%	76	0.0%
Mary	12	21.4%	16	28.6%	21	37.5%	5	8.9%	2	3.6%	0	0.0%	56	0.0%
John	13	25.0%	9	17.3%	22	42.3%	2	3.9%	3	5.7%	3	5.8%	52	0.0%
Michael	14	35.9%	6	15.4%	19	48.7%	0	0.0%	0	0.0%	0	0.0%	39	0.0%
Daniel	15	22.4%	19	28.4%	24	35.8%	6	9.0%	2	3.0%	1	1.5%	67	0.0%
Steven	30	43.5%	14	20.3%	20	29.0%	3	4.4%	1	1.5%	1	1.5%	69	0.0%
Anthony	18	26.1%	16	23.2%	25	36.2%	6	8.7%	4	5.8%	0	0.0%	69	0.0%
David	27	32.9%	26	31.7%	18	22.0%	6	7.3%	3	3.7%	2	2.4%	82	0.0%
Sarah	2	6.3%	3	9.4%	3	9.4%	9	28.1%	7	21.9%	8	25.0%	32	0.0%
Thomas	10	25.6%	11	28.2%	12	30.8%	4	10.3%	2	5.1%	0	0.0%	39	0.0%
Robert	31	41.9%	14	18.9%	22	29.7%	3	4.1%	2	2.7%	2	2.7%	74	0.0%

In the process of dealing with errors, students' behavior displayed one of two patterns. First, students often immediately had an idea of whether they agreed with the feedback when they read it. For example: "He says it's a weird sentence, but I actually don't agree" (John, think-aloud). This immediate response on feedback displays similarities with the 'bumping into errors' pattern in the feedback provision phase where students immediately recognized a text element as erroneous. Second, students often had to reread feedback remarks in order to interpret their meaning and decide whether they agreed with the feedback. For instance, in the following excerpt, Sarah processes feedback of Steven who marked the words 'self fulfilling' [sic] in her text yellow:

Marking? (...) I don't understand, 'formatted'. I don't understand why Steven marked this. Or is it because self-fulfilling should be written as one word? I don't know. Is he [he = *Microsoft Word's* spelling check] going to mark this as an error? [*types 'self-fulfilling' in her text.*] Ah, I see, I think he [he = Steven] means that it is one word. So then we can delete this comment (Sarah, think-aloud).

This process of simultaneously reading and interpreting a feedback remark displayed similarities with the 'discovering of errors' in the feedback provision phase, where students (iteratively) reread text elements in order to interpret them.

Effects of Interpersonal Factors while Processing Feedback

Similar to the feedback provision phase, only one think-aloud utterance was labeled as being related to semantic domain (6) 'Effects of interpersonal factors during the feedback processing phase'. However, in the interviews, there was a clear distinction between students that declared to be affected by their relationship with the feedback provider while processing the peer-feedback (Mary, James, Robert, John, David, Thomas, Daniel, Anthony), and students that said not or hardly to be affected by their relationship with the feedback provider while processing the peer-feedback (Linda, Michael, Steven, Sarah). Six students that said to be affected by the interpersonal relationship mentioned the perceived language skills (Mary, James, John, Michael, Sarah, Thomas), and four students (Daniel, Anthony, Thomas, Steven) mentioned perceived effort as an important factor. For example:

I think that I – I think that I keep that [that = the identity of the feedback provider] into account. Especially because I, ehm, well yes, you always have got an idea about someone. Like, well, is this person, ehm, skilled in the Dutch language, let's say (perceived language skills, Thomas, interview);

If I would think 'this person has not seriously worked on this task', than I would find it more difficult to take the criticism seriously, because I wouldn't know in what way it would have been written (perceived effort, Daniel, interview).

The other four students (Linda, Michael, Steven, Sarah) did not show any signs of being affected by the interpersonal relationship with the feedback provider. Additionally, these four students shared several other characteristics: they (a) were all members of class B; (b) made less comparisons with criteria than the eight students that said to be affected by the interpersonal relationship; (c) together formed two dyads; and (d) made less references to the other while processing feedback than almost all other students (see Table 5.6).

Table 5.6. Frequencies for codes related to semantic domain (5) Mentioning the other, during the processing of feedback.

	He/she	Name	You	Total
Linda	5	0	0	5
James	18	1	0	19
Mary	18	4	4	26
John	9	2	0	11
Michael	8	0	0	8
Daniel	7	2	0	9
Steven	5	0	0	5
Anthony	19	1	0	20
David	14	0	0	14
Sarah	1	2	0	3
Thomas	10	1	0	11
Robert	0	0	0	0
Average	9.5	1.1	0.3	10.9

Discussion

The aim of the current study was to investigate what cognitive sub-phases could be distinguished in the process of dealing with errors, and how these sub-phases may be affected by interpersonal factors during the provision and processing of peer-feedback. With a think-aloud protocol, we explored students' thoughts during the activities of peer-feedback provision and peer-feedback processing, and in semi-structured interviews students looked back on the peer-feedback activities.

Our think-aloud data showed that the order in which the sub-phases of the process of dealing with errors occurred during the provision of peer-feedback, partly deviated from the expected behavior based on the conceptualization by Aben et al. (2019). Whereas Aben et al. (2019) hypothesized a sequence of error identification, decoding, evaluation, and feedback encoding, our results showed that some sub-phases may be skipped, may follow each other so rapidly that they could not be distinguished with the current measurement method, or may occur in a different order. In the feedback provision phase, we found two patterns. In the pattern of 'bumping into errors', the identification of errors seemed to happen simultaneously with the decoding, and often any thoughts related to an evaluation phase were lacking, which was comparable to the pattern as found by Ahmadian et al. (2019). By contrast, during the pattern of 'discovering of errors', the identification of an error seemed to occur as a result of an interpreting/evaluating phase.

In the feedback processing phase, we found two patterns, that, to a certain extent, were comparable with the patterns in the feedback provision phase. In general, the patterns in the processing phase were also only partly in line with the conceptualization by Aben et al. (2019). Whereas Aben et al. (2019) hypothesized a sequence of feedback reading, decoding, evaluating, and output revising, the reading and decoding phases could not be distinguished in students' think-aloud utterances.

Although our analyses did not show relations between the types of errors and how those errors were dealt with during the processing phase (as found by Kim and Bowles, 2019), we did find evidence that the process of dealing with errors had different appearances dependent of types of errors during the provision of peer-feedback. More specifically, the pattern of 'bumping into errors' mostly resulted in the

identification of errors related to lower order concerns, whereas the pattern of 'discovering of errors' mostly resulted in the identification of errors related to higher order concerns. Hence, future research could continue to explore the relationship between the appearance of sub-phases during the provision of peer-feedback and different types of errors.

The analyses showed that the majority of students in our sample predominantly focused on errors, and comparatively less on pluses, while providing and processing peer-feedback. Whereas they declared that they identified pluses about as often as errors while providing feedback, errors resulted more often in the encoding of a feedback remark than pluses. Furthermore, during the feedback processing phase, textual revisions only occurred as a result of feedback related to errors and not as a result of feedback related to pluses. The focus on errors is not surprising. Students are primarily involved in peer-feedback activities with the aim to improve their (peers') learning (performance), rather than complimenting their (peers') current learning states (Liu & Carless, 2007). As the necessity to bridge a gap in relation to a standard is not there in the case of pluses, this likely enhances students' error-oriented focus in feedback activities (Narciss, 2017). This is also in line with Máñez et al. (2019), who found that students spend more time on errors than on pluses while processing feedback.

With respect to the influence of interpersonal factors on feedback processes, we detected differences when we compared the feedback provision and processing phase. Our findings suggest that interpersonal factors influenced students only implicitly during the feedback provision phase, as almost all of the students said in the interview not to be affected by the interpersonal relationship, whereas half of them showed before, after, or during the think-aloud activities, or in the interview to be aware of the recipient's perceived language skills and/or perceived effort. By contrast, during the feedback processing phase, interpersonal factors seemed to affect the majority of the students explicitly in the interview. Two third of the students declared in the interview taking the provider's perceived language skills and/or the perceived effort into account while processing their feedback. These results were in line with previous research illustrating that peer-feedback processing may be affected by perceptions of language skills (Berndt et al., 2018; Strijbos et al., 2010) and perceived effort (Timmers et al., 2020).

Regarding interpersonal factors, the difference between the feedback provision and feedback processing phase is remarkable. On

the one hand, it could be that students want to portray themselves as objective, and/or strive to be objective, when providing feedback. Research shows that avoiding bias is highly valued when making decisions or when assessing (Irwin & Real, 2010), which is also reflected in the wide range of attempts to compose descriptions of performance quality criteria, such as rubrics (Panadero & Jonsson, 2013). On the other hand, during feedback processing, the feedback recipient's self-interest may be too large to ignore knowledge of the interpersonal relationship that potentially conveys information about the accuracy of the provided feedback.

Strengths and Limitations

The combination of a quantitative and a qualitative analysis provided a richer account of the way students dealt with errors and the role of interpersonal factors during the provision and processing of peer-feedback than either analysis would have provided. Findings from both analyses were consistent, for example, in showing that students focus more on errors than on pluses when providing and processing feedback, and that interpersonal factors play a role in the processing of peer-feedback. Simultaneously, the analyses provided complementary information. For example, the qualitative analysis revealed that the occurrence of sub-phases, as identified in the quantitative analysis, followed different patterns. Additionally, the qualitative analysis showed that students seemed to be explicitly affected by interpersonal factors in the feedback processing phase, and implicitly in the feedback provision phase. This emphasizes the value of combining multiple data sources and analytical techniques in order to better understand the provision and processing of feedback.

Simultaneously, it is important to bear in mind potential limitations of this study. First, one could argue that social desirability played a role in students not reporting that the perceived Dutch language skills of their peer may have affected their provision of feedback. However, as students had no problems with mentioning that interpersonal factors had affected their feedback processing behavior, our data rather seemed to indicate that most students were not aware of effects of interpersonal factors on their feedback provision behavior, whereas most of them were aware of those effects during the processing of feedback.

Second, there are known limitations of think-aloud studies in general. For example, Sachs and Polio (2007) found that learners

who did not have to think aloud made more accurate revisions while processing teacher feedback on their text written than learners who did have to think aloud. In fact, students in our sample mentioned that they were unexperienced in thinking aloud, which may have affected their feedback provision and processing behavior positively—“Actually I think my mind would wander much quicker if I would not have to talk aloud” (Mary, interview)—or negatively—“Well I – yeah, it was a bit difficult, because of course you only had ehm, yeah, five seconds let's say, to think” (Anthony, interview). Especially this last remark emphasizes caution, as it shows that Anthony did not understand the rationale behind thinking aloud.

Practical Implications

Despite the limitations, the results have clear implications for educational practice. First, earlier research emphasized the crucial role of teacher instruction for optimal learning gain from peer-feedback activities (e.g., Min, 2005; Van Steendam et al., 2010). With respect to the process of dealing with errors, this implies that the explanation of criteria and the deviations from criteria, e.g. errors, may also carry a central role in instructions for peer-feedback activities. Our results showed that the students of the one class referred more often and more efficiently to criteria than the students of the other class. Although our sample was too small to investigate whether this was due to a classroom or teacher effect, it suggests that the engagement with text quality criteria is either a result of, or a prerequisite for, efficiently dealing with errors while providing and processing peer-feedback on writing performance.

Second, instructions prior to peer-feedback activities in educational settings should also aim to teach students to view their texts in a holistic manner. The students in our sample typically immediately started with reading the argumentative text (feedback provision phase) or the feedback (feedback processing phase) on a micro level. As such, they did not appear to look at their texts with a bird's-eye view, and therefore did not treat the text and feedback as coherent entities. This also became clear from the fact that most of the students identified a complete new set of errors on micro level when they reread the text when providing feedback. The development of a holistic view on texts improves the chances of identifying errors related to higher order concerns, which in turn may lead to more significant text improvement (Lerchenfeldt et al., 2019).

Third, the fact that interpersonal factors seemed to influence

feedback provision and feedback processing behavior implies that the composition of a dyad may also influence potential learning gains. Especially when processing peer-feedback, students may, for example, spend less time on processing feedback when they perceive the effort as invested into the provided peer-feedback as low. This implies that peer-feedback activities could be optimized when students as well as teachers would be aware of the potential role that interpersonal factors may play in peer-feedback provision and processing.