

Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing

Markus Berndt^{1,2} · Jan-Willem Strijbos³ · Frank Fischer⁴

Received: 17 October 2016 / Revised: 21 June 2017 / Accepted: 24 July 2017 /

Published online: 13 August 2017

© Instituto Superior de Psicologia Aplicada, Lisboa, Portugal and Springer Science+Business Media B.V. 2017

Abstract Peer-feedback efficiency might be influenced by the oftentimes voiced concern of students that they perceive their peers' competence to provide feedback as inadequate. Feedback literature also identifies mindful processing of (peer)feedback and (peer)feedback content as important for its efficiency, but lacks systematic investigation. In a 2×2 factorial design, peer-feedback content (concise general feedback [CGF] vs. elaborated specific feedback [ESF]) and competence of the sender (high vs. low) were varied. Students received a scenario containing an essay by a fictional student and fictional peer feedback, a perception questionnaire, and a text revision, distraction, and peer-feedback recall task. Eye tracking was applied to measure how written peer feedback was (re-)read, e.g., glance duration on exact words and sentences. Mindful cognitive processing was inferred from the relation between glance duration and (a) text-revision performance and (b) peer-feedback recall performance. Feedback by a high competent peer was perceived as more adequate. Compared to CGF, participants who received ESF scored higher on positive affect towards the peer feedback. No effects were found for peer-feedback content and/or sender's competence level on

✉ Markus Berndt
markus.berndt@med.uni-muenchen.de; <http://dam.klinikum.uni-muenchen.de>

Jan-Willem Strijbos
j.w.strijbos@rug.nl; <http://www.rug.nl/staff/j.w.strijbos>

Frank Fischer
frank.fischer@psy.lmu.de; http://www.psy.lmu.de/ffp_en/index.html

¹ Institute for Medical Education, University Hospital of LMU, Ziemssenstr. 1, D-80336 Munich, Germany

² Richard W. Riley College of Education and Leadership, Walden University, 100 Washington Avenue South, Suite 900, Minneapolis, MN 55401, USA

³ Department of Educational Sciences, University of Groningen, Grote Rozenstraat 3, 9712 TG Groningen, The Netherlands

⁴ Department of Psychology, Ludwig-Maximilians-Universität, Leopoldstr. 13, 80802 Munich, Germany

performance. Glance durations were negatively correlated to text-revision performance regardless of condition, although peer-feedback recall showed that a basic amount of mindful cognitive processing occurred in all conditions. Descriptive findings also hint that this processing might be dependent on an interaction between peer-feedback content and sender's competence, signifying a clear direction for future research.

Keywords Peer assessment · Peer feedback · Mindful cognitive processing · Eye tracking · Cooperative/collaborative learning

The qualitative component of peer-assessment approaches—i.e., peer feedback—has become a major research focus in online and face-to-face settings (Nelson and Schunn 2009; Prins et al. 2006). Peer feedback can be provided more frequently and more quickly than feedback by the instructor (Falchikov and Goldfinch 2000) and can foster collaboration skills, communication skills, and personal responsibility (Liu and Carless 2006; Mwalongo 2013; Sadler 2010).

Peer feedback received less attention in educational research compared to feedback in general, but it is subject to the same limitations when it comes to feedback efficiency. First of all, feedback in general does not automatically lead to positive outcomes (Hattie and Timperley 2007; Kluger and DeNisi 1996; Narciss 2008; Shute 2008). Secondly, several authors proposed *mindful cognitive processing* as important for feedback efficiency (Bangert-Drowns et al. 1991; Gielen et al. 2010; Narciss 2008; Poulos and Mahony 2008; Strijbos et al. 2010), but systematic investigations are limited. Thirdly, several studies on peer feedback in recent years revealed that interpersonal aspects might influence peer-feedback efficiency. For example, students lack confidence in their own and their peers' domain knowledge (Hanrahan and Isaacs 2001; Lin et al. 2002) and the ability to provide peer feedback (Ballantyne et al. 2002; Van Gennip et al. 2010).

The overall research question of the present study investigates the impact of peer-feedback content and sender's competence level on peer-feedback perceptions and peer-feedback application (performance), together with the role of mindful cognitive processing for peer-feedback efficiency. The following sections will address current research on the effects on performance of various feedback contents, characteristics of the feedback sender, and feedback perceptions with specific attention to (academic) writing and revision as written peer feedback allows the application of eye tracking methodology to operationalize mindful cognitive processing.

Feedback content, perception, and performance

Feedback refers to “all post-response information that is provided to a learner to inform the learner on his or her actual state of learning or performance” (Narciss 2008, p. 127). Depending on the characteristics of the feedback and the instructional setting, feedback can have a positive effect on performance but also no effect or even a negative effect (Bangert-Drowns et al. 1991; Kluger and DeNisi 1996). Feedback content, among other aspects, varies in the extent of information provided which appears related to its efficiency for performance (improvement) (Narciss and Huth 2004; Nelson and Schunn 2009).

Several perspectives on feedback content adopt a multidimensional view (Butler and Winne 1995; Hattie and Timperley 2007; Kluger and DeNisi 1996; Narciss 2008). Narciss (2008)

stresses three main components influencing feedback efficiency: function, content, and form. Narciss differentiates between the cognitive, meta-cognitive, and motivational function of feedback, which are influenced by individual factors of the recipient such as learning goals or meta-cognitive skills. Narciss' classification further distinguishes two components of feedback content: the evaluative or verification component (outcome-related information) and the informative or elaboration component (additional information) and further categorizes feedback types into (a) simple feedback types, containing outcome-related information such as "knowledge of result" or whether a task has been addressed correctly, and (b) elaborated feedback types, containing additional information such as "knowledge on how to proceed" or providing specific examples. Finally, the form of feedback specifically addresses the presentation of feedback contents in terms of timing (e.g., immediate vs. delayed), adaptivity (e.g., adaptive vs. non-adaptive), and modality (e.g., spoken vs. written).

The value of specific feedback content for performance (or performance improvement) has received much attention but the research findings are mixed. Although Kluger and DeNisi (1996) observed an overall positive impact of corrective feedback (e.g., the correct answer is A) on performance (131 papers, 607 effect sizes; $d = .41$), the effects were heterogeneous with a negative tendency in over 30% of the included studies. Only some studies support the assumption that more elaborated feedback, providing more specific information (e.g., information about mistakes or on how to proceed), has a more positive effect on performance compared to feedback providing merely an evaluation or rather general comments (e.g., information about results) (Hattie and Timperley 2007; Narciss 2008; Shute 2008; Van der Kleij et al. 2015). Although many authors called for further research on feedback content and stressed the importance of mindful (cognitive) processing for a positive impact of feedback on performance (Bangert-Drowns et al. 1991; Kluger and DeNisi 1996; Narciss 2008; Srijbos et al. 2010)—i.e., the conscious (mindful) expenditure of attention directed to the content of a feedback message might enhance cognitive processing, subsequent application, and therefore result in improved performance—research to date is very limited, with Bolzer et al. (2015) as a point in case, introducing a first operationalization of mindful cognitive processing.

In terms of feedback perception, when investigated, most studies typically address its perceived usefulness. The central role of perception in feedback processing is differentiated by Ilgen et al. (1979) as follows: feedback needs to (a) be perceived, (b) be accepted as accurate, (c) be perceived as useful, and (d) result in actual behavior change. In response to Mory (2004)—who argued for the need to identify measurable variables that can reflect internal cognitive and affective processing of learners to determine effects on feedback perception and utilization—Poulos and Mahony (2008) concluded from a qualitative study that students' perception of the feedback provider influences the credibility and hence the impact of feedback. In line with multidimensional perspectives on feedback, measures for the perceptions of feedback have also moved beyond unidimensional instruments, for example, measuring feedback utility, sensitivity, confidentiality, and retention (King et al. 2009) or more general feedback usefulness and competence support (Rakoczy et al. 2013). Although being multidimensional, both instruments mainly focus on the perceptions of teacher feedback, as well as on general feedback rather than on specific feedback content.

In the domain of peer feedback, Tsui and Ng (2000) compared teacher feedback and peer feedback in secondary schools and found that students perceived feedback from both sources as helpful, as long as it provided elaborated and specific information for revision. Although students appear to prefer elaborated peer-feedback types over general peer-feedback types, this does not automatically imply that elaborated peer feedback is valid and/or effective. For

example, Strijbos et al. (2010) showed that although elaborated specific peer feedback was perceived as more adequate by the recipient during the intervention phase of their experiment, concise general peer feedback had a more positive effect on performance. They also reported more negative affect for students who received elaborated specific feedback from a high competent peer and suggested that, under certain conditions, a more elaborated peer feedback may negatively affect performance. It seems advisable to treat general assumptions on perceptions and effects of elaborated specific and concise general peer feedback with caution (Kollar and Fischer 2010; Strijbos et al. 2010; Topping 2010). This adds to the overall mixed evidence for effects of (peer)feedback content.

Feedback source, perceptions, and performance

Besides feedback content, characteristics of the feedback source, i.e., the feedback sender, contribute to the effectiveness of feedback and are suggested to be important for feedback perception and acceptance (Greller and Herold 1975; Ilgen et al. 1979; Leung et al. 2001). A common sense assumption would be that feedback from an expert is regarded as more useful than feedback from a layperson or a peer/colleague with similar knowledge and/or status. However, it is important to take into account that expertise of a feedback sender might not necessarily be perceived as such by the recipient. This might especially be the case for feedback from a peer, given the relatively similar status of the feedback sender and receiver.

Peer feedback literature provides some clues on how a feedback sender is perceived by the recipient. In the domain of (academic) writing and revision, students prefer expert or teacher feedback (Kwok 2008; Lockhart and Ng 1993; Zhang 1995) and perceive peer feedback as helpful and welcome to improve their writing (Kwok 2008; Tsui and Ng 2000) but are simultaneously concerned about their own and their peers' ability to provide feedback (Ballantyne et al. 2002). Irrespective of these preferences, Cho and Schunn (2007) found that students improved their writing quality more when receiving feedback from multiple peers than receiving feedback from a single expert.

However, there is also evidence that perceptions of peer feedback and subsequent performance are not necessarily correlated. In a study by Kaufman and Schunn (2010), students received written peer-feedback comments on a writing assignment from fellow students which they then had to apply in a revision. Students' perceptions of this peer feedback were measured in terms of fairness, usefulness, and validity. Students most often voiced concerns about fairness, especially about the qualification and competence of the peer to provide feedback. Students' perceptions of the peer feedback—regardless whether it was positive or negative—did not correlate with revision performance. Similar findings were reported by Strijbos et al. (2010) from a laboratory study with scenarios and manipulated peer-feedback content. During the intervention, (a) feedback on an essay revision task by a high competent peer was perceived as more adequate than feedback from a low competent peer, (b) feedback from a high competent peer was perceived as most adequate, (c) elaborated specific feedback (ESF) from a high competent peer resulted in more negative affect than concise general feedback (CGF), and (d) students who received CGF performed better in the essay revision task compared to students who received ESF. However, on the posttest, students who had received manipulated feedback from a low competent peer during the intervention outperformed students who received feedback from a high competent peer. Strijbos et al. (2010) concluded that complicated interactions might exist between peer-feedback content and competence of

the feedback sender on perceptions and subsequent performance. As a possible explanation, they suggested that receiving ESF from a high competent peer might render the student passive and less inclined to apply the peer feedback and stressed the need to further investigate the role of mindful (cognitive) processing for peer-feedback application.

Mindful cognitive processing of feedback

Mindfulness has its origins in the work by Salomon and Globerson (1987) who describe “mindful engagement” as a means to improve learning outcomes. They argue that mindfulness occurs in learning settings that require active and generative processes involving mental effort. Bangert-Drowns et al. (1991) appropriated this understanding of mindfulness and applied it in feedback research arguing that “(...) intentional feedback for retrieval and application of specific knowledge appears to stimulate the correction of erroneous responses in situations where it is mindful.” (p. 213). They argue that characteristics of feedback can either diminish or stimulate mindfulness, that is, “feedback is most effective under conditions that encourage learners’ mindful reception” (Bangert-Drowns et al. 1991, p. 233); hence, they stress the importance of mindfulness in feedback operations. Mindfulness refers to how deeply the feedback has been cognitively processed and understood—thus, it means *mindful cognitive processing*. In other words, if feedback is not mindfully cognitively processed, it may not result in application and subsequently not enhance (or even decrease) performance. Although mindfulness, mindful reception, or mindful processing have been proposed as a crucial factor for feedback efficiency (Gielen et al. 2010; Kluger and DeNisi 1996; Narciss 2008; Poulos and Mahony 2008; Strijbos et al. 2010), research is scarce.

To make mindful cognitive processing more tangible in the context of written peer feedback, eye tracking methodologies are ideal to uncover this implicit measure, for example, during the reading phase when processing written peer feedback. The eye-mind hypothesis states that what a person consciously attends to is also cognitively processed (Just and Carpenter 1980) and measuring it can uncover important aspects about students’ learning processes (Knight et al. 2014). Fixations are used to calculate glance duration to reveal how much attention is allocated to a specific area (Mikkilä-Erdmann et al. 2008). Eye tracking has been applied in various research domains—e.g., reading comprehension (Rayner 1998, 2009), novice versus expert attentional skills (Van Gog and Scheiter 2010), self-monitoring while reading (Hyönä et al. 2002), and decision making (Muldner et al. 2009)—but rarely in the domain of peer feedback. To address this gap, glance duration as a measure for processing peer feedback can be connected to subsequent performance measures to infer mindful cognitive processing.

The present study, research questions, and hypotheses

The present study applied a 2×2 experimental design in the domain of academic writing in higher education to investigate the impact of peer-feedback content (concise general feedback [CGF] vs. elaborated specific feedback [ESF]) and sender’s competence level (low vs. high) on peer-feedback perceptions, text-revision performance, peer-feedback recall performance, and mindful cognitive processing. The study used a scenario-based approach to enable the systematic and controlled variation of peer-feedback content and competence level of the peer-

feedback sender. Scenarios have been shown to invoke almost identical reactions from persons, comparable to real situations (Robinson and Clore 2001). Eye tracking methodology was used to elicit data about the peer-feedback reading process and obtain more insight into mindful cognitive processing.

Research question 1: What is the effect of peer-feedback content and sender's competence level on peer-feedback perceptions? Based on the findings of Strijbos et al. (2010) and Kaufman and Schunn (2010), three hypotheses have been formulated. It was expected that elaborated specific peer feedback leads to (a) higher perceived adequacy of feedback (PAF), (b) more willingness to improve (WI), and (c) more negative affect (AF) (hypothesis 1a). Literature reveals that students prefer feedback from an expert rather than a peer; however, perceptions of peer feedback seem to be mixed. In line with Strijbos et al. (2010), it is expected that feedback from a high competent peer is perceived as (a) more adequate compared to feedback from a low competent peer, (b) leads to more willingness to improve, and (c) results in a more negative affect (hypothesis 1b). They also found evidence for a multivariate interaction between peer-feedback content and sender's competence level on perceptions (PAF, WI, and AF). Hence, combining hypothesis 1a and 1b, it is expected that ESF by a high competent peer is perceived as the most adequate, but also leads to a more negative affect (hypothesis 1c).

Research question 2: What is the effect of peer-feedback content and sender's competence level on text-revision performance and peer-feedback recall performance? Based on research comparing teacher feedback versus peer feedback, and findings that students prefer ESF to CGF (Strijbos et al. 2010; Tsui and Ng 2000), it is expected that students who received ESF from a high competent peer will outperform all other research conditions for text-revision performance (hypothesis 2a) and will furthermore show a better peer-feedback recall performance (hypothesis 2b).

Research question 3: What is the role of mindful cognitive processing of peer feedback? In the present study, mindful cognitive processing follows the operationalization of Bolzer et al. (2015) as the correlation between glance duration on peer feedback with (a) text-revision performance and (b) peer-feedback recall performance. Feedback provided by a competent sender is preferred by feedback recipients (Lockhart and Ng 1993; Zhang 1995). It therefore was hypothesized that feedback from a high competent peer leads to longer glance duration than feedback from a low competent peer—regardless of feedback content (hypothesis 3a). Based on the eye-mind hypothesis (Just and Carpenter 1980), glance duration is regarded as an indicator for cognitive processing. Accordingly, it was expected that there would be a positive correlation between glance duration on peer feedback and (a) text-revision performance (hypothesis 3b) and (b) peer-feedback recall performance (hypothesis 3c).

Method

Study design and participants

The present study investigated the impact of peer-feedback content and sender's competence on peer-feedback perceptions, text-revision performance, peer-feedback recall performance, and mindful cognitive processing. Elements of the experimental procedure and tasks reported by Strijbos et al. (2010) were re-used with permission for the present study. Participants

received a scenario in which a fictional student received either CGF or ESF by a fictional peer who was either high (H) or low (L) competent.

The participants were 45 native speaker psychology students (10 male, 35 female). Their age ranged from 18 to 50 years ($M = 25.13$, $SD = 7.23$); participation was voluntary. The students were randomly assigned to one of four conditions: CGF-H ($N = 10$), CGF-L ($N = 12$), ESF-H ($N = 11$), and ESF-L ($N = 12$).

Materials

Task and peer-feedback scenarios

The study was situated in the context of academic writing and revision. The participants were asked to (a) read an essay written by a fictional student, (b) read feedback for this essay provided by a fictional peer, (c) revise the essay according to the peer feedback, and (d) recall the peer feedback after a distraction phase. Prior to reading the essay, participants received information on four main text comprehension criteria on which the peer feedback was based (simplicity, structure, conciseness, and stimulation), each consisting of several subcriteria (example for structure: "Subsections and structured, logical and correct order, well-organized (sections and indentation) distinguish major and minor issues, story line"). This was preferred over peer feedback on the content level for which the individual knowledge of each participant would have to be taken into account. The task was re-used with permission from the study by Strijbos et al. (2010, see Appendix A from their study for more details). The essay by a fictional student was an adaption of the text "Foundations of sexuality" from Langer, Schulz von Thun and Tausch (1999, pp. 44–56). The essay was prepared with 31 typical errors (i.e., long sentences, repetition of expressions, missing explanations) of which two served as examples for the participant on how to revise the essay.

The peer feedback was also based on the text comprehension criteria. The participants received either CGF or ESF coming from either a high or a low competent peer. CGF provided a short knowledge of result feedback (example for "conciseness": "The text is rather long and long-winded"). ESF additionally provided type and position of errors as well as information on how to improve the essay (example for conciseness: "Both with respect to language usage and the content long-winded. The 'Meanwhile' section is nearly superfluous, the examples too extensive" (see for more details Strijbos et al. 2010, Appendix B). The peer feedback was presented on one screen, divided into four paragraphs. Each paragraph consisted of one of the four text comprehension criteria with the respective peer feedback underneath (Fig. 1). Before reading the peer feedback, participants twice received on-screen information about the peer-feedback sender: "On the following screen you will receive feedback from a peer who, thus far, has achieved [high/low] performances in academic writing."

Manipulation check and pilot testing

After the reading phase of the peer feedback, a manipulation check was implemented. The participant was asked to recall the previously provided information whether the feedback was provided by a peer "who, thus far, has achieved [high/low] performances in academic writing." An incorrect answer would result in the exclusion of the participant after the trial.

Prior to the main data collection, the experimental materials were piloted on three psychology students to ensure that the materials were suitable to elicit complex eye movement data.

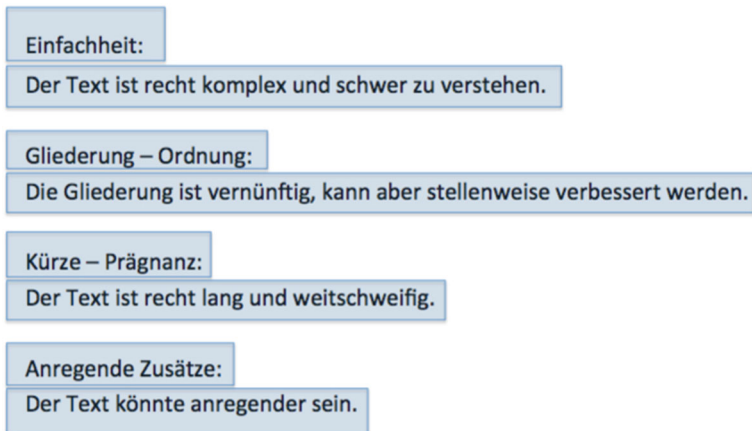


Fig. 1 Text comprehension criteria, concise general feedback (CGF), and AOIs (*boxes*)

After the first two pilot students, the materials were altered in a fashion that allowed for the best calibration process of the eye tracker to achieve maximum accuracy. The third pilot student served as a confirmatory pilot of the final materials. These three students were not included in the main sample.

Measures

Peer feedback perceptions questionnaire

Perceptions of the peer feedback were measured with the questionnaire by Strijbos et al. (2010). The questionnaire consisted of four scales with three items each: (a) Fairness (e.g., “I would consider this feedback fair”; $\alpha = 0.90$); (b) Usefulness (e.g., “This feedback would provide me a lot of support”; $\alpha = 0.94$); (c) Acceptance (e.g., “I would dispute this feedback”; $\alpha = 0.84$); and (d) Willingness to Improve (e.g., “I would be willing to improve my performance”; $\alpha = 0.84$); and one scale with six items: (e) Affect (e.g., “I would feel offended if I would have received this feedback”; $\alpha = 0.77$). Combined, the fairness, usefulness, and acceptance scales constitute “Perceived Adequacy of Feedback” ($\alpha = 0.93$). All items were measured on a 10-cm visual analogue scale from 0 (*fully disagree*) to 10 (*fully agree*). Negatively phrased items were recoded so that all scales measured positive dimensions.

Eye tracking and glance duration

Eye movements were recorded with a head-mounted Dikablis 25-Hz eye tracker by Ergoneers LLC with a 4.2-mm objective. Before the start of each trial, the first page of text was used to calibrate the eye tracker to line-by-line accuracy and to measure reading speed of the participants. Directly before the screen with the peer feedback, recording of eye movements commenced. Once the reading of the peer feedback ended, the head-mounted eye tracker was removed. After each trial, glance duration was computed on predefined areas of interest (AOIs). In combination with visual markers, the Dikablis software allowed for dynamic AOIs which automatically adapt to movements of the head. This allowed the participant to move more naturally while performing the task without affecting the accuracy of the measurement.

There were eight AOIs, one for (a) each of the four peer-feedback criteria and one for (b) each corresponding peer-feedback content; all AOIs were presented together on one screen: see example for CGF in Fig. 1. This also allowed for accumulation of the data into (a) overall glance duration on criteria, (b) overall glance duration on peer feedback, and (c) overall glance duration on criteria and peer feedback. Following the example of Bednarik and Tukiainen (2008, p. 100), the “proportional fixation time (PFT)” for each of the eight AOIs was computed as a ratio of the glance time on an area to the overall glance time and accumulated to overall proportional glance duration.

Text-revision performance and peer-feedback recall performance

The essay contained 31 errors referring to the four text comprehension criteria. Two of the 31 errors were provided as examples together with an instruction to illustrate how to revise the essay. Participants had to identify errors in the essay and make a revision proposal. Only if both criteria were fulfilled, the participant was awarded with a score point. Time on task was measured during text revision; however, in the instruction, the participants were informed that they have as much time for the task as desired. Text-revision performance was operationalized as the ratio of (a) the number of correctly identified errors and correct text-revision proposals (numerator) relative to (b) the time needed for text revision (denominator). Time was a relevant factor in order to adequately capture performance, e.g., if a participant corrected 10 errors in less time than another participant, performance was rated higher.

After revising the essay, participants were asked to recall the peer feedback in as much detail as possible. To ensure proper recall, the students were distracted beforehand with an unrelated task: “Solve as many Sudoku puzzles as you can in ten minutes.” Peer-feedback recall performance was operationalized as the ratio of the number of correctly recalled aspects of the peer feedback to the total number of aspects contained in the peer feedback.

Mindful cognitive processing

The correlation between glance durations on the peer feedback and text-revision performance as well as peer-feedback recall performance serves as an indicator to infer mindful cognitive processing of the peer feedback. Eye tracking data provides glance durations containing information about how the peer feedback has been read. If the participant makes corrections in the text-revision task which have also been suggested in the peer feedback, mindful cognitive processing can be inferred. This is depicted by a positive correlation between glance duration and text-revision performance. This operationalization also holds true for all other possible outcomes, for instance, if there is a high text-revision performance without mindful cognitive processing of the peer feedback (i.e., short glance duration on the peer feedback), this is depicted by a low or negative correlation. This operationalization of mindful cognitive processing is essentially not about glance duration per se and text-revision performance per se but about what happened in between; what has been read, mindfully processed, and applied or recalled, depicted by the correlation between measures.

Procedure

The experiment consisted of two phases. At the start of the first phase (on-screen), the eye tracker was mounted on the head of the participant and calibrated. Next, the participant was

instructed to start the on-screen scenario by pushing a button on a keyboard and, from there on, read every screen and push the button autonomously. Participants were first presented with a brief text explaining the four text comprehension criteria, followed by a revised text from a fictional student. After reading the text, participants received on-screen information about the competence of the peer-feedback sender, as well as the information that they will be asked to apply the feedback later on. Next, the participants received feedback from a fictional peer. The information about the competence of the peer-feedback sender (high vs. low) and peer-feedback content (CGF vs. ESF) were systematically varied. During the entire first phase, the participants wore the eye tracker.

In the second phase (pen and paper), participants first answered a manipulation check question prompting them to recall the previously provided information about the competence of the peer-feedback sender. Then, participants proceeded to rate the peer feedback with the perception questionnaire. Next, they had to apply the peer feedback to a second revision of the text. After the text revision, participants were distracted for 10 min by solving Sudoku puzzles and then asked to recall, with as much detail as possible, the previously read peer feedback including the text comprehension criteria.

Each experimental trial was conducted in the same laboratory under identical conditions. The experiment lasted approximately 75 min on average, depending on calibration of the eye tracker and time needed for text revision and peer-feedback recall.

Results

Distribution assumptions were checked through standardized skewness and kurtosis for all dependent variables. Skewness was inside the -3 to $+3$ range for all variables (Tabachnick and Fidell 2013). Kurtosis was outside this range for overall glance duration on criteria (7.79) and text-revision time (7.07), but after exclusion of two participants with large deviations (> 3 SDs), kurtosis for these variables was also inside the range. Overall, seven univariate outliers were detected and checked for influence on mean, skewness, and kurtosis. In six cases, the difference to the next most extreme non-outlying value was marginal and was therefore retained. In one case for the variable text-revision time, the value was adapted to one unit above the nearest highest score (Tabachnick and Fidell 2013). This adaption did not have an effect on any of the dependent measures. No structural bivariate or multivariate outliers were found. There were no missing values. The manipulation check question was correctly answered by all 45 participants.

Research question 1: Impact of peer-feedback content and sender's competence level on perceptions

A 2×2 MANOVA with peer-feedback content and sender's competence level was calculated for WI, AF, and PAF. Levene's test indicated homogeneity of variance (WI = .48, AF = .70, and PAF = .18). No multivariate interaction between peer-feedback content and sender's competence level was found, Pillai's trace = .01, $F(3, 41) = 0.126, p = .094$, but the multivariate main effect for peer-feedback content (Pillai's trace = .25, $F(3, 41) = 4.27, p = .011, \eta^2 = .25$) and sender's competence level (Pillai's trace = .21, $F(3, 41) = 3.39, p = .027, \eta^2 = .21$) were significant. Table 1 provides the means and standard deviations for WI, AF, and PAF by group.

Table 1 Mean and standard deviation for willingness to improve (WI), affect (AF), and perceived adequacy of feedback (PAF) as a function of group

	CGF-H		CGF-L		ESF-H		ESF-L	
	(N = 10)		(N = 12)		(N = 11)		(N = 12)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
WI	7.91	1.80	6.30	1.80	7.26	1.82	6.02	1.80
AF	4.29	1.55	4.05	1.52	5.42	1.53	5.08	1.52
PAF	5.78	2.02	4.08	2.01	6.86	2.02	4.91	2.01

Note: CGF-H = concise general feedback by a high competent peer; CGF-L = concise general feedback by a low competent peer; ESF-H = elaborated specific feedback by a high competent peer; ESF-L = elaborated specific feedback by a low competent peer

Follow-up univariate analyses for the main effects were performed with the Bonferroni correction set at alpha level $< .05/3$. A significant main effect for peer-feedback content on AF was found, $F(1, 44) = 5.49$, $p = .024$, $\eta^2 = .12$, with ESF leading to more positive affect ($M_{\text{ESF}} = 5.24$; $SD = 1.67$) than CGF ($M_{\text{CGF}} = 4.16$; $SD = 1.31$). There was also a significant main effect for sender's competence level on WI, $F(1, 44) = 6.92$, $p = .012$, $\eta^2 = .14$, with feedback from a high competent peer ($M_{\text{high}} = 7.57$; $SD = 1.60$) leading to higher willingness to improve than from a low competent peer ($M_{\text{low}} = 6.16$; $SD = 1.93$) and for sender's competence level on PAF, $F(1, 44) = 9.25$, $p = .004$, $\eta^2 = .18$, with feedback from a high competent peer ($M_{\text{high}} = 6.35$; $SD = 1.79$) leading to higher perceived adequacy of feedback than from a low competent peer ($M_{\text{low}} = 4.49$; $SD = 2.20$). Further post hoc analyses showed a significant moderator effect of PAF on sender's competence level for WI, $F(1, 43) = 12.62$, $p = .001$, $\eta^2 = .23$. PAF was significantly positively correlated to WI, $r(45) = .56$, $p < .01$.

Research question 2: Impact of peer-feedback content and sender's competence level on text-revision performance and peer-feedback recall performance

A significant positive but moderate correlation between text-revision performance and peer-feedback recall performance was found, $r(45) = .31$, $p < .05$. Next, a 2×2 MANOVA with peer-feedback content and sender's competence level was performed for text-revision performance and peer-feedback recall performance. The analysis yielded neither a multivariate interaction effect (Pillai's trace = .009, $F(3, 41) = 0.175$, $p = .840$) nor a multivariate main effect for peer-feedback content (Pillai's trace = .064, $F(3, 41) = 0.138$, $p = .264$) or sender's competence level (Pillai's trace = .035, $F(3, 41) = 0.726$, $p = .490$). Univariate analyses showed no main effects for the different groups. Descriptives showed a non-significant pattern ($p = .101$) towards better peer-feedback recall in the CGF group: $M_{\text{CGF}} = 0.37$ ($SD = 0.26$), $M_{\text{ESF}} = 0.26$ ($SD = 0.16$), $d = .52$ (medium effect). Descriptively, peer-feedback recall declined with increasing amount of information and declining competence of the sender (see Fig. 2).

Research question 3: Mindful cognitive processing of peer feedback

As ESF is per se longer than CGF, it is only natural that it produces longer glance durations and tests for differences are therefore redundant. The impact of sender's competence level on overall glance duration was investigated with two separate ANOVAs for CGF and ESF. No

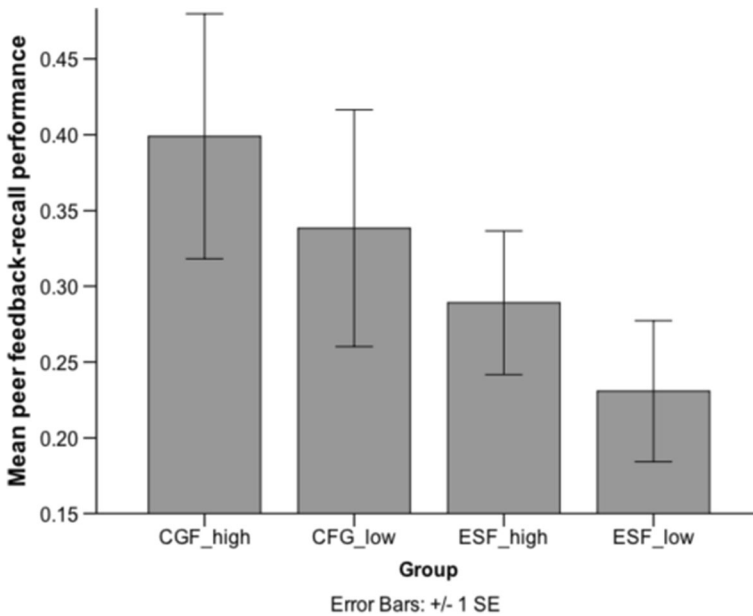


Fig. 2 Means of peer-feedback recall performance per group

effects were found for either group, $F(1, 21) = 0.09, p = .766$ (CGF), $F(1, 22) = 0.01, p = .926$ (ESF). Subsequently, more in-depth ANOVAs with (a) glance duration on only criteria and (b) glance duration on only peer-feedback content also yielded no effects.

The relationship between overall glance duration and text-revision performance was first examined across conditions (ESF, CGF, high and low). No relationship was found between overall glance duration on peer feedback and text-revision performance. More in-depth analyses for (a) glance duration on only criteria and (b) glance duration on only peer-feedback content revealed a significant negative correlation of medium strength between glance duration on only criteria and text-revision performance, $r(45) = -.39, p < .01$. Next, the relationships in the ESF and the CGF groups were analyzed separately. In the CGF group, a significant negative correlation of medium strength was found between overall glance duration and text-revision performance, $r(22) = -.54, p < .05$, and for glance duration on only peer-feedback content and text-revision performance, $r(22) = -.52, p < .05$. In the ESF group, a significant negative correlation of medium strength was found between overall glance duration and text-revision performance, $r(23) = -.51, p < .05$, for glance duration on only criteria and text-revision performance, $r(23) = -.43, p < .05$, and for glance duration on only peer-feedback content and text-revision performance, $r(23) = -.51, p < .05$. Finally, the relationships in the competence-high and competence-low groups were analyzed separately. In the competence-high group, a significant negative correlation of medium strength was found between glance duration on only criteria and text-revision performance, $r(21) = -.44, p < .05$. No relationships were found in the competence-low group.

The relationship between overall glance duration and peer-feedback recall performance was first examined across conditions (ESF, CGF, high and low). No relationship was found between overall glance duration on peer feedback and peer-feedback recall performance. More in-depth analyses were performed for (a) glance duration on only criteria and (b) glance duration on only peer-feedback content, but no correlations were found. Subsequent separate

analyses for the ESF and the CGF groups revealed no significant relationships. Finally, separate analyses were performed for the competence-high and the competence-low groups, which revealed no relationships.

Discussion

The aim of the present study was to investigate the impact of peer-feedback content and sender's competence level on peer-feedback perceptions, text-revision performance, peer-feedback recall performance, and mindful cognitive processing. In this section, the results with regard to the formulated hypotheses will be discussed.

Effect of peer-feedback content and sender's competence level

Students who received ESF only reported a more positive affect (AF) but no higher willingness to improve (WI) or higher perceived adequacy of feedback (PAF). Therefore, hypothesis 1a was rejected. Feedback was perceived as more adequate when provided by a high competent peer which in turn increased WI. Likewise, feedback from a high competent peer resulted in more WI and hypothesis 1b was therefore partially confirmed. No interactions were observed between peer-feedback content and competence level of the peer-feedback sender (hypothesis 1c was not confirmed). Hence, in line with Strijbos et al. (2010), students' perception of adequacy of the peer feedback was higher when knowing that the feedback was provided by a high competent peer which in turn increased their WI. This relates well to reports that students in general prefer feedback from an expert (Lockhart and Ng 1993; Zhang 1995). However, Strijbos et al. (2010) reported a negative influence of ESF by a high competent peer on AF and also interactions between peer-feedback content and sender's competence level on AF, which we did not observe in our data. In the present study, ESF resulted in a more positive affect. These findings suggest that students' concern with their peers' ability to provide feedback in writing and revision (Ballantyne et al. 2002) might be influenced in either direction when the student receives information about the competence of the peer-feedback sender. As we manipulated the competence of the peer-feedback sender with a label we attached to the peer feedback and checked whether the participants recalled this label correctly, this might not be adequate to capture the participants' own perspective but merely provided some information for the participant about the competence of the sender. It appears worthwhile to conduct further research to obtain more insight into possible interaction effects between peer-feedback content and sender characteristics on peer-feedback perceptions.

Neither significant interaction effects nor significant main effects were found for peer-feedback content and sender's competence level on text-revision performance and peer-feedback recall performance, possibly due to power issues. Hypotheses 2a and 2b were rejected. In retrospect, information about participants' baseline skill in academic writing and revision would have been an interesting mediator variable and could be considered in future studies. In terms of peer-feedback recall, descriptives showed a non-significant pattern towards the best recall for students that received CGF from a high competent sender. Peer-feedback recall performance descriptively decreased with (a) increasing information in the peer feedback and (b) lower competence of the peer-feedback sender—i.e., more elaborated (and longer) peer feedback and a lower competence level of the peer-feedback sender descriptively resulted in weaker peer-feedback recall performance. This hints at a possible interaction effect between

peer-feedback content and sender's competence level on peer-feedback recall performance and also indicates that the amount of information provided plays an important role. Future research should examine this in more detail.

Mindful cognitive processing of peer feedback

As glance duration on ESF is naturally longer than glance duration on CGF, only data was analyzed for differences between the competence-high and competence-low groups. It appears that the peer-feedback reading process is not influenced by information about the competence of the peer-feedback sender (hypothesis 3a was not confirmed).

Mindful cognitive processing was operationalized as the correlation between glance duration on peer feedback with (a) text-revision performance and (b) peer-feedback recall performance. In terms of text-revision performance, no significant correlations were found for overall glance duration on peer feedback across conditions (ESF, CGF, high and low). An in-depth analysis revealed significant negative correlations in both the CGF and ESF conditions between (a) overall glance duration and text-revision performance, (b) glance duration on content and text-revision performance, and (c) glance duration on criteria and text-revision performance (only in ESF condition). With regard to sender's competence, a significant negative correlation was found between glance duration on criteria and text-revision performance in the competence-high condition. Hypothesis 3b was therefore rejected. Counterintuitively, longer glance duration on peer feedback appears to be related to lower text-revision performance—i.e., spending more time on reading the peer feedback does not necessarily lead to better application of the peer feedback in the text-revision task. It should be noted that we do not have information about why a participant showed longer glance durations on the peer feedback. A possible explanation could be deeper, more intense processing, as well as difficulties in cognitive processing. Longer glance durations merely are an indicator for longer cognitive processing. In the light of cognitive load theory, an explanation could be that the task of simultaneously processing semantic properties of both peer feedback and corresponding essay text qualifies as a task with high element interactivity. The degree of element interactivity is the most important aspect that determines the complexity of information (Kalyuga 2011) and intrinsic cognitive load imposed on the participant (Sweller 1994; Sweller et al. 1998). Very high (or very low) intrinsic cognitive load is widely recognized to reduce performance and inhibit learning.

In terms of peer-feedback recall performance, no significant correlations were found for overall glance duration on peer feedback across conditions (ESF, CGF, high and low); thus, hypothesis 3c was rejected. Nevertheless, students were able to recall certain aspects of the peer feedback which signifies that mindful cognitive processing did occur. Combined with the descriptive pattern that peer-feedback recall appears to decrease with the amount of information provided, a tentative conclusion is that students might have suffered from an overload of information. This underlines Shute's (2008) perspective to "provide elaborated feedback in small enough pieces so that it is not overwhelming or discarded" (p. 177). Moreover, when the feedback was provided by a low competent peer, in both the text-revision and peer-feedback recall task, a descriptive pattern for lower performances was observed. One possible explanation could be that feedback by a low competent peer is perceived as less adequate and might reduce mindful cognitive processing. Taking into account reports that the ability of the sender can play a role for feedback acceptance (Kali and Ronen 2008; Lin et al. 2001; Strijbos et al. 2010), there might also exist a similar effect for mindful cognitive processing, that is, feedback

provided by a low competent peer is more readily regarded as irrelevant and thus less mindfully cognitively processed. As illustrated by Fig. 2, when feedback was provided by a low competent peer, peer-feedback recall performance was lower even within peer-feedback content groups.

It should be noted that this study was one of the first to define, operationalize, and systematically investigate *mindful cognitive processing* in the context of peer feedback. Further research on mindful cognitive processing—i.e., operationalization and effects—is much needed.

Methodological limitations

The medium-strength positive correlation between text-revision performance and peer-feedback recall performance indicates that both are adequate measures for performance, each covering different aspects. However, the operationalization of mindful cognitive processing through the relationship between glance duration on peer feedback and both performance measures served to gain more insight into underlying cognitive processes when written peer feedback is perceived and applied. To identify and include further eye tracking measures to confirm inferences made from glance durations seems to be an important task for future research. A possible measure could be the number of gaze shifts between essay text and peer feedback as an attempt of the student to relate them (Bolzer et al. 2015). Similar transitions have been used by Mason et al. (2013) as an indicator of integrative effort between verbal and graphical representations.

One common problem in research applying eye tracking methodologies is the typically small sample size. For instance, in expertise research—depending on the domain—an average sample consists of 11 experts, 10 intermediates, and 12 novices (Gegenfurtner et al. 2011). Although data from an above-average sample size has been elicited in the present study (as compared to typical sample sizes in eye tracking studies), the sample size could have been larger to avoid possible influence from sampling error and raise statistical power. Larger samples to confirm, reject, or modify existing findings seem to be an important challenge for future research.

Practical implications

The present study revealed that teachers and peers should not assume that elaborated peer feedback—whether in an online or face-to-face setting—automatically leads to better application and performance. It seems important that elaborated peer feedback—specifically the large amount of information it contains—is provided in small enough doses so as not to inhibit mindful cognitive processing by the receiver. Teachers should therefore not only prompt students to provide elaborated peer feedback, but also not to “overprovide.” Information about the competence level of the peer-feedback sender appears important for perceived adequacy of the peer feedback, but not necessarily for its application. Therefore, teachers can opt to include students who are perceived as low competent in peer-feedback settings. Nevertheless, the perception of peer feedback is important and bears a huge (de-) motivational potential for students. Together with the general preference of students for elaborated peer feedback, it seems advisable to foster a certain level of competence in the students, to ensure a positive perception of adequacy of the peer feedback and to reduce negative affect.

Acknowledgements The authors would like to thank Tamara van Gog and Frans Prins for their comments and suggestions to the present study.

References

- Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education*, 27(5), 427–441.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213–238.
- Bednarik, R., & Tukiainen, M. (2008). Temporal eye-tracking data: Evolution of debugging strategies with multiple representations. *Proceedings of the 2008 symposium on eye tracking research & applications* (pp. 99–102). New York, NY: ACM.
- Bolzer, M., Strijbos, J. W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning*, 31, 422–434.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, 65(3), 245–281.
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the disciplines: a web-based reciprocal peer review system. *Computers & Education*, 48(3), 409–426.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287–322.
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educational Psychological Review*, 23, 523–552.
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, 20(4), 304–315.
- Greller, M. M., & Herold, D. M. (1975). Sources of feedback: a preliminary investigation. *Organizational Behavior and Human Decision Processes*, 13(2), 244–256.
- Hanrahan, J., & Isaacs, G. (2001). Assessing self- and peer assessment: the students' views. *Higher Education Research and Development*, 20, 53–70.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hyyönä, J., Lorch Jr., R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: evidence from eye fixation patterns. *Journal of Educational Psychology*, 94(1), 44–55.
- Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), 349–371.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, 87, 329–355.
- Kali, Y., & Ronen, M. (2008). Assessing the assessors: added value in web-based multi-cycle peer assessment in higher education. *Research and Practice in Technology Enhanced Learning*, 3(1), 3–32.
- Kalyuga, S. (2011). Cognitive load theory: how many types of load does it really need? *Educational Psychology Review*, 23(1), 1–19.
- Kaufman, J. H., & Schunn, C. D. (2010). Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science*, 39(3), 387–406.
- King, P. E., Schrodt, P., & Weisel, J. J. (2009). The instructional feedback orientation scale: conceptualizing and validating a new measure for assessing perceptions of instructional feedback. *Communication Education*, 34(2), 235–261.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.
- Knight, B. A., Horsley, M., & Eliot, M. (2014). Eye tracking and the learning system: an overview. In M. Horsley, M. Eliot, B. A. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research* (pp. 281–285). Cham: Springer International.
- Kollar, I., & Fischer, F. (2010). Commentary—Peer assessment as collaborative learning: a cognitive perspective. *Learning and Instruction*, 20(4), 344–348.
- Kwok, L. (2008). Students' perceptions of peer evaluation and teachers' role in seminar discussions. *Electronic Journal of Foreign Language Teaching*, 5(1), 84–97.
- Langer, L., Schulz von Thun, F., & Tausch, R. (1999). *Sich verständlich ausdrücken* [Expressing oneself comprehensively]. München, Germany: Ernst Reinhardt Verlag.

- Leung, K., Su, S., & Morris, M. W. (2001). When is criticism not constructive? The roles of fairness perceptions and dispositional attributions in employee acceptance of critical supervisory feedback. *Human Relations, 54*, 1155–1187.
- Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: feedback for students with various thinking-styles. *Journal of Computer Assisted Learning, 17*, 420–432.
- Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2002). Student attitudes toward networked peer assessment: case studies of undergraduate students and senior high school students. *International Journal of Instructional Media, 29*, 241–256.
- Liu, N.-F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education, 11*(3), 279–290.
- Lockhart, C., & Ng, P. (1993). How useful is peer response? *Perspectives, 5*, 17–29.
- Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education, 81*(3), 356–384.
- Mikkilä-Erdmann, M., Penttinen, M., Anto, E., & Olkinuora, E. (2008). Constructing mental models during learning from science text. Eye tracking methodology meets conceptual change. In D. Ifenthaler, P. Pirnay-Dummer, & J. M. Spector (Eds.), *Understanding models for learning and instruction: Essays in honor of Norbert M. Seel* (pp. 63–79). New York: Springer.
- Mory, E. H. (2004). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed., pp. 745–783). New York, NY: Erlbaum.
- Muldner, K., Christopherson, R., Atkinson, R., & Burleson, W. (2009). Investigating the utility of eye-tracking information on affect and reasoning for user modeling. In G.-J. Houben, G. McCalla, F. Pianesi, & M. Zancanaro (Eds.), *User modeling, adaptation, and personalization* (pp. 138–149). Berlin: Springer.
- Mwalongo, A. I. (2013). Peer feedback: its quality and students' perception as a peer learning tool in asynchronous discussion forums. *International Journal of Evaluation and Research in Education, 2*(2), 69–77.
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. Van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 125–143). Mahwah, NJ: Erlbaum.
- Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multi-media learning. In H. M. Niegemann, D. Leutner, & R. Brünken (Eds.), *Instructional design for multimedia learning* (pp. 181–195). Münster: Waxmann.
- Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: how different types of peer feedback affect writing performance. *Instructional Science, 37*(4), 375–401.
- Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: the students' perspective. *Assessment & Evaluation in Higher Education, 33*, 143–154.
- Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for general practitioners in training: quality, styles, and preferences. *Advances in Health Sciences Education, 11*, 289–303.
- Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: mediated by students' perception, moderated by goal orientation. *Learning and Instruction, 27*, 63–73.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.
- 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.
- Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35*(5), 535–550.
- Salomon, G., & Globerson, T. (1987). Skill may not be enough: the role of mindfulness in learning and transfer. *International Journal of Educational Research, 11*, 623–637.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.
- Strijbos, J. W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: are they critical for feedback perceptions and efficiency? *Learning and Instruction, 20*(4), 291–303.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Allyn & Bacon.

- Topping, K. J. (2010). Commentary—Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, 20(4), 339–343.
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Learning*, 9(2), 147–170.
- Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: a meta-analysis. *Review of Educational Research*, 85(4), 475–511.
- Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: the role of interpersonal variables and conceptions. *Learning and Instruction*, 20(4), 280–290.
- Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, 20(2), 95–99.
- Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing*, 4(3), 209–222.

Markus Berndt. Institut für Didaktik und Ausbildungsforschung in der Medizin, Klinikum der Universität München (LMU), Ziemssenstr. 1, D-80336 Munich, Germany. Tel.: +49-89- 4400-57208. Fax.: +49-89-4400-57202. Email: markus.berndt@med.uni-muenchen.de, Website: <http://dam.klinikum.uni-muenchen.de>

Current themes of research:

- a) Peer assessment and peer feedback in higher education
- b) Scientific reasoning and argumentation in higher education
- c) Diagnostic competencies and clinical reasoning in medical education

Relevant publications in the field of Psychology of Education:

- Bolzer, M.,** Strijbos, J. W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: the role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning*, 31(5), 422–434. doi:10.1111/jcal.12091
- Schworm, S., & **Bolzer, M.** (2014). Learning with video-based examples—are you sure you do not need help?. *Journal of Computer Assisted Learning*, 30(6), 546–558. doi:10.1111/jcal.12063

Jan-Willem Strijbos. University of Groningen, Department of Educational Sciences, Grote Rozenstraat 3, 9712 TG Groningen, the Netherlands. Tel.: +31-50-363-6658. Email: j.w.strijbos@rug.nl; Website: <http://www.rug.nl/staff/j.w.strijbos/>

Current themes of research:

- a) Peer assessment and peer feedback
- b) (Computer-supported) collaborative learning
- c) Formative assessment

Most relevant publications in the field of Psychology of Education:

- Bolzer, M., Strijbos, J. W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: the role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning*, 31(5), 422–434. doi:10.1111/jcal.12091
- Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on peer assessment: effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, 39(4), 195–203.
- Strijbos, J. W. (2011). Assessment of (computer-supported) collaborative learning. *IEEE Transactions on Learning Technologies*, 4(1), 59–73.
- Strijbos, J. W., Narciss, S., & Dünnebie, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: are they critical for feedback perceptions and efficiency? *Learning and Instruction*, 20(4), 291–303.

Strijbos, J. W., & Sluijsmans, D. M. A. (2010). Unravelling peer assessment: methodological, functional, and conceptual developments. *Learning and Instruction, 20*(4), 265–269.

Frank Fischer. LMU Munich, Department of Psychology and Munich Center of the Learning Sciences, Leopoldstr. 13, 80802 Munich, Germany. Tel.: +49-89-2180-5146. Email: frank.fischer@psy.lmu.de; Website: http://www.psy.lmu.de/ffp_en/index.html

Current themes of research:

- a) Computer-supported collaborative learning
- b) Fostering diagnostic competences in simulation-based learning environments
- c) Scientific reasoning and argumentation
- d) Executive functions in complex learning environments

Most relevant publications in the field of Psychology of Education:

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). [Toward a script theory of guidance in computer-supported collaborative learning](#). *Educational Psychologist, 48*(1), 56–66.

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., et al. (2014). [Scientific reasoning and argumentation: advancing an interdisciplinary research agenda in education](#). *Frontline Learning Research, 2*(3), 28–45.

Schwaighofer, M., Bühner, M., & Fischer, F. (in press). Executive functions as moderators of the worked example effect: when shifting is more important than working memory capacity. *Journal of Educational Psychology*.