

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/65378> holds various files of this Leiden University dissertation.

Author: Huisman, B.A.

Title: Peer feedback on academic writing : effects on performance and the role of task-design

Issue Date: 2018-09-12



General discussion

7

Within a broader context of institutional initiatives to increase study success and decrease student attrition across bachelor programs, this thesis focused on the design of formative peer feedback tasks and the effects thereof on students' academic writing performance. In particular, this thesis investigated to what extent formative peer feedback impacts higher education students' academic writing performance and how particular aspects of peer feedback task-design affect students' writing performance. An important aim was to amalgamate both scientific and practical value. In terms of its practical value for higher education teaching staff, this thesis focused on relatively controllable aspects of peer feedback task-design. In terms of its contribution to the peer feedback literature, this thesis aimed to contribute to the limited number of studies adopting relatively controlled research designs. To reiterate, formative peer feedback was defined as 'all task-related information that a learner communicates to a peer of similar status which can be used to modify his or her thinking for the purpose of learning'. Hence, formative peer feedback in this thesis could include information ranging from simple scores to elaborate comments, generally encompassing both 'peer assessment' and 'peer feedback' insofar as these reflect different practices in the literature.

Main Findings per Chapter

Chapter 2. The impact of formative peer feedback on writing performance: A meta-analysis

This study described a meta-analysis of quantitative studies reporting on the impact of formative peer feedback on higher education students' academic writing performance. Two sets of research questions were addressed. First, the synthesized effects of peer feedback on academic writing were compared both to a no-feedback baseline and to two oftentimes feasible alternatives: self-assessment and feedback from teaching staff. Second, we explored the moderating role of two variables that are simultaneously important in the design of peer feedback tasks (Gielen et al., 2011) and that were perceived as controllable by higher education teaching staff. Specifically, the number of peers that a student engaged with during peer feedback and the nature of the peer feedback (scores,

comments, or both) were assessed with respect to their moderating role in explaining students' writing improvement after peer feedback. Results indicated that the effects of peer feedback on academic writing performance tend to be larger than that of either no feedback or self-assessment, whereas the effects of peer- and teacher feedback do not appear to differ significantly. Results of the moderation analyses indicated that a combination of both peer comments and –scores tends to have a larger effect on writing performance than either peer comments or –scores alone. In contrast, the number of peers with whom a student engaged during peer feedback did not moderate subsequent writing performance, although there appeared to be a trend suggesting that engagement with multiple peers is more beneficial than engagement with a single peer. These results suggested that higher education teaching staff design their peer feedback tasks in such a way that students provide both comments and scores to each other's writing. A notable limitation of this meta-analysis concerned the limited number of studies that proved eligible for inclusion. We argued that this limitation signals an important observation regarding the current body of literature on the topic, namely that there is a need for more well-controlled, (quasi-)experimental studies into peer feedback on writing in HE.

Chapter 3. Students' ability match, the nature of peer feedback and writing performance

The literature remains inconclusive with regard to how students are optimally matched during peer feedback. Meanwhile, the increasing availability and user-friendliness of (often web-based) applications increases the options and opportunity for higher education teaching staff with respect to the design and implementation of peer feedback tasks. The third chapter described a quasi-experimental study that explored the relation between students' ability match, the nature of the peer feedback and subsequent performance in the context of an authentic academic writing task. Contrasting two different perspectives on student ability matching from the literature, students were assigned to a dyad together with either a similar-ability or different-ability peer. Results indicated that students in similar-ability and different-ability dyads do not differ

in the extent to which they improve their writing. In addition, neither dyad composition nor students' individual ability significantly related to the nature of the peer feedback they provided, although a trend suggested that content-related peer feedback may be more prevalent within similar-ability dyads consisting of two high ability students. Finally, the nature of the received peer feedback did not relate to the subsequent writing performance of either lower or higher ability authors. This last finding suggested that relatively high or low ability authors do not benefit differently from peer feedback on either content-, structure- or style-related aspects of their writing. In summary, students' ability match, the nature of the peer feedback and academic writing performance were not significantly related. Some remarks were in place, however. For one, this study did not take into account students' perceptions with respect to the peer feedback they received (cf. Hanrahan & Isaacs, 2001; Strijbos et al., 2010). Also, this study did not disentangle the separate effects that providing and receiving peer feedback could have on students writing performance (e.g., Lundstrom & Baker, 2009; Cho & MacArthur, 2011). Consequently, how students' peer feedback perceptions related to their subsequent writing performance remained a question for future research, and the reported effects of peer feedback on students' writing reflected the combined effects of both providing and receiving peer feedback. Finally, the students in this study were relatively similar in terms of gender, age and educational background, which could mean that the differentiation between higher and lower ability students was one within a restricted range. This may explain why only non-significant trends were found and, at the same time, suggests that these trends could be considered informative; if these patterns emerged within a sample that is relatively homogeneous in terms of students' ability, they may become more salient as heterogeneity increases, such as in open online educational contexts.

Chapter 4. Students' ability match and academic writing performance in a MOOC

The number of massive open online courses (MOOCs) and participants therein has rapidly expanded over recent years. This has increased the relevance of issues

such as formal recognition of learning and accreditation, and raises the question how to optimally scale up the assessment of complex, open-ended assignments. In MOOCs, the answer to this oftentimes involves peer assessment in the form of both comments and grades between multiple peers. Following up on the on-campus study described in chapter three, this study conducted a first exploration of how participants' ability match during peer feedback relates to their writing performance. The subjects in this study were 565 participants in a Leiden University MOOC on Terrorism and Counterterrorism. They were categorized as either relatively high, intermediate or low performers based on available performance metrics prior to the first essay assignment. Post hoc analyses were conducted to assess the relation between participants' own performance level, that of their reviewing peers', and participants' subsequent increase in writing performance between essays. Results indicated that the average performance level of the assessing peers positively relates to participants' subsequent increase in writing performance. A closer look revealed an interaction between participants' own performance level and that of their peer assessors. Specifically, peer assessors' average performance level only related to the writing performance of intermediate and higher performing participants, not to that of relatively low performing participants. Effect sizes were small, however. Possible interpretations of these results could be that participants' relatively low, intermediate and high performance level also reflects their ability to utilize peer feedback. Alternatively, or complimentary, these findings indicated that peer feedback quality could be improved. Certain limitations needed to be considered when interpreting the results of this exploratory study. For example, the lack of background information on participants left open the question of what the exact mechanisms were through which the performance level of participants and that of their peer assessors interacted. Similarly, it also remained a question for future research what the role of peer feedback quality and quantity is in explaining MOOC participants' writing performance. Still, this study provided a first insight in how the match between MOOC participants and their peer assessors related to participants' writing performance, contributing to the knowledge base on how to optimally design peer assessment tasks in MOOCs.

Chapter 5. Students' peer feedback role, peer feedback perceptions and writing performance

Research on the mechanisms involved in the peer feedback process has focused on elements with respect to both the provision of peer feedback and the reception thereof. On the one hand, a learning-by-reviewing rationale suggests that providing peer feedback stimulates students' active problem detection, – diagnosis, and subsequent contemplation of revision strategies (e.g., Flower et al., 1986; Patchan & Schunn, 2015). On the other hand, receiving (peer) feedback provides students with information on the gap between current performance, goal performance, and how to close that gap (e.g., Nicol & Macfarlae-Dick, 2006; Hattie & Timperley, 2007). However, how these different aspects of the peer feedback process compare in terms of their impact on students' academic writing performance largely remains an open question. This study compared the impact of students' feedback role – being either the provider or receiver of peer feedback – on their subsequent performance increase in the context of an academic writing task. Additionally, the relations between the nature of the received peer feedback, students' perceptions thereof, and performance increase were investigated. Results indicated that feedback providers and receivers improved their writing to similar degrees, and that explanatory peer feedback mostly influenced the extent to which students perceived the received peer feedback as adequate. However, no direct relations were found between students' perceptions of the received peer feedback and their subsequent increase in writing performance. As a no-feedback control-group was not ethically feasible, one limitation to bear in mind was that this study only referred to the *relative* effect that providing or receiving peer feedback has on students' writing performance. Nevertheless, these results were believed to elucidate two findings that are informative for higher education researchers and –teaching staff alike. First, for students that are unfamiliar with the peer feedback process, students' role may at first be that of peer feedback provider only. Initially withholding received peer feedback could avoid issues such as students' distrust of their peers' feedback quality, while the exercise of providing peer feedback would still be likely to be beneficial to students' writing performance. Second,

there was the finding that explanatory peer feedback most strongly related to students' perceptions of adequacy, which was argued to be important with respect to students' more general support for – and engagement in – the peer feedback process. Recognizing the importance of students' support for peer feedback, which is likely to be shaped by multiple peer feedback experiences, this finding highlighted the importance of emphasizing the role of explanations in peer feedback training and instruction.

Chapter 6. A questionnaire to assess students' beliefs about peer feedback

Research on peer feedback may focus on different types of outcomes, including performance, perceptions of learning and beliefs about assessment (van Gennip, Segers, & Tillema, 2009). In case of students' beliefs about peer feedback, research appears to vary both in terms of the instruments used and in terms of reported outcomes (see van Zundert, Sluijsmans, & van Merriënboer, 2010, for an overview). Chapter six described the development of the Beliefs about Peer Feedback Questionnaire (PBFQ). This questionnaire served a dual aim. For one, it aimed to contribute to the alignment and, consequently, comparability of research findings. Simultaneously, it aimed to provide a practical instrument for higher education teaching staff to monitor how their teaching practice influences students' peer feedback beliefs. Four subscales were conceptualized based on an initial set of eleven items, tested in a separate exploratory study ($N = 219$) and a confirmatory study ($N = 121$). The a priori conceptualized model with four scales was consistently found to best fit the data. These four scales related to: (1) students' valuation of peer feedback as an instructional method, (2) students' confidence in the quality of the peer feedback they provide to a peer, (3) students' confidence in the quality of the peer feedback they receive from a peer and (4) the extent to which students regard peer feedback as an important skill. Individual scale reliabilities ranged between $\alpha = .67$ and $\alpha = .82$ across both samples. Although additional testing across other institutes and disciplines would be desirable to further confirm the external validity of the PBFQ, a first successful step in this respect was made by purposefully conducting the confirmatory analyses on a different group of students. Hence, we believe

that this study provides a comprehensive instrument to assess students' peer feedback beliefs. Additionally, we believe that the concise nature of the BPFQ makes it a practically applicable instrument, both for higher education teachers that want to conduct research within their own teaching practice as for researchers that aim to monitor the development of students' peer feedback beliefs (e.g., longitudinal designs).

Conclusion

The overarching research question in this thesis addressed two related issues: the available evidence for the impact of formative peer feedback on students' academic writing performance and the role of specific peer feedback task-design aspects in explaining writing performance. Regarding the available evidence, the findings presented in the meta-analysis (chapter two) may be considered as somewhat surprising. Despite the continuous development of a framework on the most relevant aspects of peer feedback task-design (cf. Topping, 1998; van den Berg, Admiraal, & Pilot, 2006b; van Gennip, Segers, & Tillema, 2009; Gielen, Dochy, & Onghena, 2011), and despite repeated calls by scholars for more systematic and controlled studies into the effects of peer feedback (e.g., Topping, 1998, 2010; Strijbos & Sluijsmans, 2010), only 16 studies with *some type of* reference group were eligible for inclusion. This clearly signals a gap in the current peer feedback literature that is worth noticing and that merits the attention of future research. Keeping in mind the limitations that this restricted number of studies imposes, the meta-analysis provided provisional evidence suggesting that engagement in peer feedback positively affects writing performance compared to a no-feedback baseline and compared to self-assessment, whereas no difference was found when compared to the effects of feedback from teaching staff.

Regarding the role that specific aspects of peer feedback task-design have in contributing to students' writing performance, chapters two through five covered four different task-design aspects that tend to be relatively controllable for HE teaching staff. First, the available evidence does not indicate that engaging in feedback with multiple peers contributes to larger writing improvements

than engaging in feedback with a single peer. However, this should probably be considered as a tentative finding because of the combination of a) the limited number of studies included in the meta-analysis, b) the non-significant pattern pointing towards larger writing improvements when multiple peers are involved and c) the theoretical arguments that would support a multiple-peer hypothesis. Second, the nature of peer feedback appears to matter in terms of students' perceptions of peer feedback adequacy and can relate to their writing performance depending on how the nature of peer feedback is defined. With respect to students' perceptions, the presence of explanatory peer feedback (i.e. peers' comments that explain or justify the peer feedback) best predicts how adequate the peer feedback is perceived to be. With respect to writing performance, a combination of peer scores and –comments appears to be more effective than either scores or comments alone. Focusing specifically on peer comments, however, the nature of those comments does not appear to be directly related to students' writing performance. Third, matching students based on prior performance appears to yield different results in different contexts. In particular, student matching within an on-campus course did not relate to subsequent writing performance, whereas the match between participants in a MOOC did positively relate to writing performance for the majority of them. Fourth, providing and receiving peer feedback appear to have a similar impact on students' own writing performance.

In addition to the role of these specific task-design aspects, over the course of the current thesis it became increasingly clear that students' support for peer feedback as an integral part of their education is pivotal. This insight developed based on various experiences during this thesis, including in-class observations and informal talks with students and with teaching staff. Recognizing the potential influence of students' beliefs about peer feedback on their related perceptions and behavior, and recognizing that these beliefs are likely to be shaped by multiple experiences over time, the Beliefs about Peer Feedback Questionnaire was developed. This questionnaire was developed to be both concise and thematically comprehensive, facilitating longitudinal inquiries into students' peer feedback beliefs.

Taken together, the current thesis furthers our knowledge on a) the available evidence for the impact of formative peer feedback on writing performance, b) how students' ability match and feedback role as either peer feedback provider or –receiver relate to writing performance, and c) the relations between the nature of the peer feedback and students' perceptions thereof. Provided that future research confirms these findings, several implications follow. For one, formative peer feedback positively affects higher education students' academic writing performance. In fact, it appears to be equally effective as formative feedback from teaching staff, which implies that higher education teaching staff can indeed regard peer feedback as an instructional method that simultaneously benefits students' learning and is logistically efficient in the case of large student-to-teacher ratios. However, it is far from self-evident that students think of teacher feedback as equally effective as peer feedback (cf. Liu & Carless, 2006; van Zundert et al., 2010; McConlogue, 2015), which could lead them to appraise the feedback they receive differently. There are understandable reasons for students' differing expectations. For example, teachers and students are not identical in terms of domain-specific knowledge and experience, which may result in feedback comments with a different focus (e.g., Cho & MacArthur, 2010). The current thesis found first-year students' peer feedback to be predominantly focused on aspects of writing style, and less on content- or structure-related aspects, although a trend suggested that higher performing dyads provide somewhat more content-related peer feedback than mixed or lower performing dyads. This could imply that peer feedback should be considered as a complimentary feedback source to teacher feedback during the first year(s) of higher education programs, and that it may be regarded as increasingly comparable to teacher feedback as students acquire more domain-specific knowledge and experience. The reported nature of first-year students' peer feedback also suggests that they should be trained and guided in providing good quality peer feedback. In particular, the findings presented in this thesis suggest that, from students' own perspective, such guidance training should emphasize the role of explanatory peer feedback comments.

These findings could be embedded in a broader framework of peer feedback training-design, which may simultaneously serve as a basis for comparing future studies that assess the effects of peer feedback training. Such a framework should include at least two complimentary aspects. On the one hand, it should include the knowledge and skills in which students are trained, including students' mastery of assessment criteria, students' ability to provide and utilize constructive peer feedback and their ability to make valid judgements of a peers' performance. On the other hand, it should include the extent to which the peer feedback training actively engages students in the process, which could for example range from simple instruction, to dialogue, to (guided) practice.

Naturally, there are strengths and weaknesses as a consequence of the methodological choices that were made and with respect to the practical implications and considerations.

Methodological reflections

The primary focus on students' academic writing performance naturally involves certain assumptions, such as that increasing grades on a writing task reflect the increase of a student's writing skills. Although it is not undebated what grades reflect (e.g., Knight, 2002), they generally remain the widest available proxy to students' learning in higher education. In addition, student performance as measured by grades aligns with Leiden University's initiatives focusing on study success. An alternative focus may be the critical appraisal of peer feedback by students (cf. Strijbos, Narciss, & Dünnebier, 2010; Gielen et al., 2011; Sadler, 2010). This is a potentially fruitful area for future research, the importance of which also became increasingly clear over the course of the current thesis. In both formal and informal conversations which are not included in this thesis, it became increasingly salient that students can have multiple motives for engaging with the received peer feedback or not. They could, for example, be skeptical with regard to their peers' legitimacy as assessor, or they were inclined to only engage in learning activities that were strictly required for passing a course. Such conversations highlighted the fact that feedback can only be expected to contribute to a student's learning when he or she mindfully engages with

it (e.g., Handley, Price, & Millar, 2011). From a more cognitive psychological perspective, this process of critical appraisal could be regarded as the bottleneck between on the one hand task- and learner characteristics and, on the other hand learning outcomes. In particular, it may be interesting to assess the interaction between specific learner characteristics (e.g., current proficiency, motivation, or task-related beliefs) and task characteristics (e.g., aspects of peer feedback task-design, the nature of peer feedback) in explaining critical appraisal. Some recent research has started to investigate such relations using, for example, eye-tracking (Bolzer, Strijbos, & Fischer, 2015), which may be a particularly promising approach to inferring a learner's mindful processing of (peer) feedback.

Another methodological reflection seems in place with respect to the focus of this thesis on the controllability of the task-design aspects. In this thesis, the notion of controllability was aligned with planned behavior theory (Ajzen, 1991; Ajzen & Fishbein, 2005). Six higher education teachers were interviewed and – using the variables from Gielen, Dochy, & Onghena's (2011) framework – performed a card-sorting task to assess which aspects of peer feedback task-design they perceived as controllable in their own teaching practice (chapter 2). Borrowing from planned behavior theory, in which perceived behavioral control is a key antecedent to intentions and behavior, teachers' perceived controllability was considered as a relevant indicator for the practical applicability of these variables. The interviewed teachers were asked to reason based upon their own specific teaching practice and experience and to weigh in practical constraints (e.g., available time and resources) when determining controllability. In other words, perceived controllability referred to teachers' perceptions of the extent to which they *could* adapt these variables. However, controllability did not directly incorporate the extent to which the teachers *would* adapt these design-aspects of peer feedback tasks. More elaborate research into the practicality (cf. Doyle & Ponder, 1977) of the different aspects of peer feedback task-design would therefore be valuable; this could provide more conclusive evidence with respect to what higher education teachers perceive they *could* and *would* change to the peer feedback tasks in their educational practice, and how these perceptions may vary between disciplines or institutes. For the six interviewed

teachers in chapter two, for example, somewhat differing patterns were found between teachers in research-intensive universities and universities of applied sciences. Possibly, these preliminary differences related to typical variations in teaching environments between these two – in the Dutch context commonly distinguished – types of universities.

A final methodological reflection concerns the absence of a no-feedback control-group in the empirical studies in this thesis. In general, studies including such a no-feedback control-group are scarce and therefore much needed to further our knowledge on *the extent to which* formative peer feedback impacts HE students' writing performance. Within the specific educational contexts of the empirical studies in this thesis, however, such control-groups were considered ethically unfeasible. Consequently, these empirical studies can only draw conclusions with regard to *relative* writing improvement for differently matched students or for students fulfilling different roles during peer feedback. The most salient objection to a no-feedback control-group that was encountered during informal conversations with both teaching staff and educational researchers, was that the omission of an existing beneficial learning element in a course is considered unethical. It seems that controlled interventions may be perceived as more acceptable in educational contexts where peer feedback was not implemented before, i.e. when peer feedback is an *additional* element instead of being *omitted as an existing* element within a course. Alternatively or complimentary, future studies into peer feedback may draw inspiration from research designs that are more common in other research domains, such as delayed treatment designs. Insofar as educational contexts would allow for such variations in research design, these may alleviate ethical concerns for quasi-experimental peer feedback studies with authentic academic writing tasks.

Practical implications and considerations

The findings of this thesis regarding the specific aspects of peer feedback task-design can be considered informative for both higher education teachers and educational advisors. Regarding the nature of the peer feedback, formative peer feedback had the most impact on students' writing performance if it included

both scores and comments. This suggests that teaching staff can optimize the impact of a peer feedback task by designing it accordingly, provided that they communicate the formative nature of the scores to their students upfront. Also, peer feedback containing explanations for accompanying value judgments or suggested revisions was particularly important for students to perceive the peer feedback as adequate. Hence, higher education teachers can optimize students' perceptions of adequacy by emphasizing the role of explanations in peer feedback. This can be achieved by systematically scaffolding the peer feedback that students provide on each other's work, for example through a combination of student training, task-instructions and rubrics to guide the peer feedback process (cf. Gielen and de Wever, 2015). In the long-term, this can positively influence students' beliefs about the importance of peer feedback as well as their confidence in themselves and their peers.

Regarding student ability matching, somewhat different conclusions were drawn in an on-campus and an online context. Irrespective of the reported significance values, however, effect sizes tended to be small in both contexts. In other words, if there is an effect of students' ability match on their subsequent increase in writing performance, this effect is probably small. This suggests that the ability-matching of higher education students should not be perceived as a top priority and that *if* ability-matching is contemplated when designing a peer feedback task, it should be weighed against the efforts required for implementation. In cases where time-constraints are salient, this renders random student matching an ethically defensible choice. When student matching becomes more or less effortless in the context of (web-based) platforms that facilitate the peer feedback process, student matching becomes a feasible option to attain these relatively small effects on students learning gains.

Regarding higher education students' feedback role, peer feedback providers and –receivers improved their academic writing to similar degrees. This knowledge should be comforting to higher education teaching staff in multiple ways. For example, when a student unintentionally does not receive feedback from a peer, the act of having provided peer feedback can still be expected to contribute to his or her own learning. This is especially relevant when group

sizes increase and, consequently, teachers' control over the implementation of the peer feedback process decreases. This knowledge is also helpful in situations where students are unfamiliar with the peer feedback process and/or are (still) skeptical with respect to their peers' feedback. In such situations, issues like students' distrust in each other's peer feedback quality may be circumvented by initially withholding the provided peer feedback from the intended receivers. As students' confidence in, and support for peer feedback as an instructional method increases with practice and experience, peer feedback may be made fully reciprocal by disclosing the peer feedback to the assessesees.

Whenever the preparation of students for a future career is considered as a broad central aim of a higher education curriculum, peer feedback should be regarded as an important learning goal in itself (e.g., Cowan, 2010; Liu & Carless, 2006; Sadler, 2010; Sluijsmans et al., 2004). Students' support for peer feedback is pivotal with respect to their engagement in the peer feedback process and the learning gains that can be expected. It therefore seems particularly worthwhile to explore how to cultivate a classroom culture where peer feedback is the norm and to investigate how students' peer feedback beliefs and skills develop over time. Possibly, higher education teaching staff could set the norm early on by informing and preparing students with respect to the peer feedback process from the very beginning of their studies. Ideally, then, peer feedback would be formative, guided in terms of the nature of the peer feedback, and gradually expanded to more elaborate and complex tasks that students believe are within the range of their developing expertise. In this light, the current thesis facilitates evidence-based decisions with respect to the design aspects of peer feedback tasks, including the nature of the peer feedback, student matching, and students' feedback role. In addition, the Beliefs about Peer Feedback Questionnaire can be instrumental in the systematic, long-term monitoring of students' peer feedback beliefs and the aspects of task-design that influence those beliefs.

In the specific context of Leiden University, the findings of this thesis are also informative for educational advisors and policy makers. Educational advisors can incorporate these findings into their training of and advice to both beginning and senior teaching staff, helping them develop peer feedback

tasks that both positively affect students' performance and that are optimally supported by the students. Educational policy makers can weigh the findings and arguments in this thesis into their decisions that affect the teaching staff's capacity for designing peer feedback tasks. By stimulating teacher training programs and by supporting the availability of peer feedback software packages, for example, educational policy makers can facilitate higher education teachers in designing effective peer feedback tasks that help students to improve their academic writing skills.