Effects of training peer tutors in content knowledge versus tutoring skills on giving feedback to help tutees' complex tasks

Effects of training peer tutors in content knowledge versus tutoring skills on giving feedback to help tutees' complex tasks

Ya Ping (Amy) Hsiao*, Francis Brouns, Jan van Bruggen, and Peter B. Sloep

Welten Institute, Open University of the Netherlands

Ya Ping (Amy) Hsiao*

E-mail: amy.hsiao@ou.nl. Telephone: +31-45-5762797. Fax: +31-45-5762800.

Acknowledgements

We would like to thank Mr. Frank Neiss, Ms Maartje van de Bovenkamp-Janssen, Mr. Rens Afman, Ms Els van Bevervoorde, Mr. Roel Verstappen, Dr. Slavi Stoyanov, and Dr. Kamakshi Rajagopal for their help during the experiment.

Abstract

This study aims to investigate the effectiveness of training tutors in content knowledge of a particular domain versus training them in tutoring skills of pedagogical knowledge when tutoring on a complex tutee task. Forty-seven tutor-tutee pairs of fourth year secondary school students were created and assigned to one of two treatments. Twenty-two tutors received training in content knowledge and the other twenty-five tutors in tutoring skills. Tutors formulated written feedback immediately after the training. Tutees first interpreted the tutor feedback and then used it to revise their research questions. The results showed that tutors trained in tutoring skills formulated more effective feedback than tutors trained in content knowledge. In addition, tutees helped by tutoring-skills tutors found the feedback more motivating than those helped by content- knowledge tutors. However, no differences were found in tutee performance on revision. The findings are discussed in terms of the set-up of this study and implications for improving the effectiveness of peer tutoring.

Keywords

Peer tutoring; tutor training; content knowledge; tutoring skills; complex tasks; peer feedback

Effects of training peer tutors in content knowledge versus tutoring skills on giving feedback to help tutees' complex tasks

Introduction

Peer tutoring is an instructional method with "people from similar social groupings who are not professional teachers helping each other to learn and learning themselves by teaching" (Topping, 1996, p.322). There is widespread agreement that peer tutoring is an effective method that benefits both tutors and tutees: on the one hand, tutors benefit from tutoring (learning twice); on the other hand, tutees benefit from their tutor's help (Falchikov, 2001; King, 1997; Topping, 1996, p.322). Whether tutees benefit from their tutor may be influenced by several aspects, two of which are the type of tutee task and the kind of help the tutor provides. These two aspects are actually interwoven. For rote learning tasks that emphasize remembering, peer tutoring may not be effective, because tutees might not need tutor help to accomplish the task. Even when peer tutoring is applied, the tutoring task is mainly to retrieve content knowledge in a particular domain and then transmit it to the tutee. Therefore, recent development on peer tutoring has shifted focus from rote learning towards complex learning (Falchikov, 2001; King, 1997, 1998). For complex tasks that emphasize higher order cognitive skills, tutees need to first understand how the facts and concepts relate with each other and then use that knowledge to find solutions. When working on complex tasks, it is likely that tutees need tutor help. To help tutees deal with ~~the relationships between facts and concepts needed for solving~~ complex tasks, peer tutors need to perform more advanced tutoring tasks than merely retrieving and transmitting knowledge. Examples of these include giving appropriate feedback for tutee responses, providing explanations for incorrect responses, and motivating tutees to maintain the level of effort needed to complete the task. Such tutoring tasks not only require tutors to *apply* the content knowledge

~~in~~to a particular domain but also pedagogical knowledge in tutoring skills (i.e., teaching techniques) (Falchikov, 2001; Lepper & Woolverton, 2002). Based on the hierarchical model of cognitive skills, performing higher order cognitive skills (e.g., applying) pre-supposes having acquired lower order cognitive skills (e.g., remembering and understanding)(Anderson & Krathwohl, 2001).

However, for most settings in which peer tutoring takes place, students who act as peer tutors are often each other's equals and they have neither sufficient content knowledge in a particular domain nor pedagogical knowledge of tutoring skills (Chi, Siler, & Jeong, 2004; King, 1998). When peer tutoring is applied ~~in~~to complex tasks, tutors without sufficient content knowledge or tutoring skills cannot successfully perform their tutoring tasks and tutees cannot benefit from the tutoring process. In turn, this will negatively impact the effectiveness of peer tutoring. Thus, to prevent this impact it is necessary to support peer tutors with lower cognitive skills (e.g., remembering and understanding) which they need to perform higher cognitive skills (e.g., applying) required by complex tasks.

Many tutoring studies suggest ~~to~~ training students prior to performing their tutoring tasks (Falchikov, 1995, 2001; Jenkins & Jenkins, 1987; King, 1991, 1994; King, Staffieri, & Adelgais, 1998; Nath & Ross, 2001; Topping, 1996). Though the effects of some training programs have been studied (Fuchs, Fuchs, Bentz, Phillips, & Hamlett, 1994; King, 1991, 1994; King et al., 1998; Nath & Ross, 2001), little attention has been paid to the relationship between the type of tutoring tasks most suited to the tutee task and the type of tutor training. In particular, we wonder which type of training can help peer tutors perform their tutoring tasks more effectively and which tutees can benefit more from tutor help when working on complex tasks. Since *none* of the tutoring studies directly compared the effectiveness of training peer tutors in content knowledge

versus tutoring skills, inferences can only be made over separate studies that investigated the effectiveness of either tutors with content knowledge or tutors with tutoring skills in different contexts. The studies we found have either examined the effects of tutors with *existing* content knowledge or of tutors' tutoring skills enhanced through prior training programs on tutee performance.

A number of tutoring studies regarded staff tutors or senior student tutors as content experts. They investigated the tutoring process for tutees who worked on complex tasks such as discussing solutions to a complex problem. Two studies found that tutees who learned with such content experts (i.e., staff tutors) performed significantly better on knowledge tests than those who learned with non-content experts (Davis, Nairn, Paine, Anderson, & Oh, 1992; Schmidt, Van der Arend, Moust, Kokx, & Boon, 1993). In contrast, two other studies found that tutees tutored by non-content experts (i.e., student tutors), performed *equally* well on a knowledge test as those tutored by content experts (i.e., staff tutors) (De Grave, De Volder, Gijselaers, & Damoiseaux, 1990; Moust & Schmidt, 1994). In addition, no difference was found in performance on a knowledge test between tutees supported by senior student tutors differing in knowledge levels (De Grave et al., 1990).

So, can tutors with content knowledge effectively help tutee learning on complex tasks? There is no unequivocal answer to this question. First, in the reported studies, tutors also possessed tutoring skills either because the studies used staff tutors or because tutors received a training that focused on tutoring skills. Therefore, the better tutee learning performance may have resulted from the combination of both tutors' content knowledge and tutoring skills. Second, although the studies mentioned above applied tutoring for complex tasks, they assessed knowledge acquisition only. Altogether these studies show that it is difficult to measure the

effect of content knowledge on learning performance in isolation. Does a similar conclusion hold for attempts to measure the effects of tutoring skills?

Unlike studies of content knowledge, in which staff and senior students acted as tutors, tutoring skills were mostly investigated in situations in which *true peers,* who are each other's equal, act as both tutors and tutees to perform *reciprocal* peer tutoring. These studies have focused on giving a training in advance to make sure that students would demonstrate enough tutoring skills to elicit desirable social interactions and to trigger cognitive processes that contribute to learning (Fuchs et al., 1994; King, 1991, 1994; King et al., 1998; Nath & Ross, 2001). The reciprocal feature of these studies meant that training was given to *all* students, who acted both as tutee and as tutor interchangeably.

These studies looked at the effect of ~~the~~ training on ~~the~~ tutoring behavior and analyzed the interaction discourse recorded from the tutoring process. They showed that after the training students demonstrated tutoring skills such as asking different types of questions (Fuchs et al., 1994; King, 1991, 1994; King et al., 1998), giving more elaborated explanations (King, 1991), as well as demonstrating cooperative and communication skills (Nath & Ross, 2001). Some of these studies have shown that students who received training in tutoring skills on higher cognitive processing (i.e., asking thought-provoking questions), outperformed those who received training on knowledge telling or those who did not receive any training of tutoring skills at all (King, 1991, 1994; King et al., 1998). This was borne out by diverse tests that measured learners' problem solving skills (King, 1991), comprehension and knowledge construction (i.e., making inferences and integrating materials) (King, 1994; King et al., 1998).

Based on these findings, it seems justified to conclude that tutoring skills are likely to facilitate the tutoring process, but the reciprocal nature of these studies makes it difficult to

conclude so unequivocally. All students of these studies have acted as both tutor and tutee and this required multiple tasks (i.e., everyone should work on at least two tasks to take on both the role of tutor and tutee). Working on multiple tasks may have introduced practice effects. Moreover, since all students received training in tutoring skills, this might ~~have~~ already have resulted in additional cognitive benefits to students' ~~their~~ own learning (e.g., they could ask themselves thought-provoking questions). Therefore, in addition to tutors' tutoring skills, both practices and additional cognitive benefits might have contributed to learning performance.

To summarize, the effects of tutors' content knowledge or training skills on the performance of complex tutee tasks are as yet inconclusive. Moreover, as discussed, there have been no earlier studies directly comparing the effectiveness of tutors trained in content knowledge to tutors trained in tutoring skills. To fill in this gap, the main research question of this study ~~will address,~~ is which type of training will make peer tutors perform more effective tutoring tasks on a complex task. We will investigate this question in the domain of formulating research questions. The training focuses on one type of tutoring task: giving feedback. To find out which tutees can benefit more from tutor help, ~~W~~we first examine tutor feedback performance to determine which type of training results in more effective tutor feedback. Since effectiveness of peer tutoring also depends on whether tutees benefit from tutor help (Topping, 1996), we then study tutee performance on research questions to determine which type of tutor's feedback is more effective on helping tutees to revise research questions. In addition, whether tutees attend to feedback (Nelson & Schunn, 2009; Strijbos, Narciss, & Dünnebier, 2010) and whether tutor feedback has stimulated tutees to maintain the needed level of effort to revise (Narciss, 2008; Narciss & Huth, 2006) are important determinants of tutee performance on

research questions. Thus, we also check tutee interpretations of tutor feedback prior to revising research questions and their motivation to revise after revision.

Method

Design

This study followed a between-subjects design with types of training as the independent variable: tutors with content knowledge (CK tutors) trained in domain specific knowledge of formulating research questions whereas tutors with tutoring skills (TS tutors) trained in pedagogical knowledge of giving feedback. To verify the relative effects of tutors' content knowledge versus tutoring skills, this study is to avoid three weaknesses found in the literature. First, to avoid mixing effects (i.e., from a combination of content knowledge and tutoring skills), either tutors' content knowledge or training skills were trained before peer tutoring started, but not both. Second, to avoid ambiguous effects from reciprocity, students in this study took on the fixed role of being either a tutor or a tutee; we trained tutors only. And third, instead of measuring knowledge acquisition, we focused on tutee performance on the task of formulating research questions, which is considered a highly complex task for secondary school students (Van der Schee, 2001).

Participants and settings

Participants were students in the fourth year of secondary education in one Dutch school: 118 students had been asked to participate, but due to an unforeseen extra-curricular event on the experimental days only 94 students actually participated (43 boys, 51 girls). Each of the 94 participants was randomly given the role of either tutor or tutee., ~~and 47~~ Forty-seven tutor-tutee

pairs were created based on random match and then each pair is randomly assigned to the treatment: 22 for the CK and 25 for the TS tutoring treatment. Participants who received help from CK tutors are called *CK tutees* whereas those who received help from TS tutors are called *TS tutees*. This study was a compulsory activity of their Natural Sciences trajectory emphasizing training in research skills. Prior to this study, all participants received one lesson taught by their teachers: formulating research questions in four steps. Each tutee had formulated a set of draft research question consisting of main and sub-questions, an example of which is shown in Table 1.

Materials

Tutor training materials

The training materials that enhanced tutors' content knowledge started with a short recap of the four steps of how to formulate a research question: orientation, limiting scope, formulating and checking against criteria. Particular attention was paid to how a good research question should be formulated based on general SMART criteria (i.e., specific, measurable, attainable, relevant and time-bound) as well as to specific criteria for research questions regarding content, structure, and consistency among main- and sub-questions. Each criterion was elaborated in definitions with good and bad examples.

The training materials that enhanced tutors' tutoring skills started with two short videos about giving and receiving feedback[1] and how to give effective feedback[2]. Particular attention was paid to how to formulate feedback in three steps: first indicate what has been done well,

[1] http://www.youtube.com/watch?v=d_Tsq7qvgW0

[2] http://www.youtube.com/watch?feature=player_embedded&v=H2eap4_TZMo

follow by what needs to be improved, and finally conclude with general positive comments. Each step was elaborated in rules of thumbs with good and bad examples.

Tutee task and tutee interpretations of tutor feedback

The tutee task was to use tutor feedback to improve their draft research questions. Since feedback left unattended cannot be effective, we introduced an intermediate step prior to tutees' revising research questions to raise a mindful reception of tutor feedback, (Nelson & Schunn, 2009; Strijbos et al., 2010). Tutees had to interpret tutor feedback by answering three open questions: i) Which feedback aspects do you find difficult to understand? ii) Which feedback aspects are you uncertain about? and iii) Which aspects do you still need your tutor to give you further feedback on?

Tutee motivation to revise

A Likert scale with eight items ranging from one (strongly disagree) to five (strongly agree) was designed to measure tutees' motivation to revise their research question based on tutor feedback (see Table 5).

Pre-measures

A prior knowledge test consisting of fifteen multiple-choice questions was administered to measure students' existing knowledge in formulating research questions. Because students' general tutoring skills and specific feedback skills might also influence the effects of the treatments, two Likert scales ranging from one (strongly disagree) to five (strongly agree) were used to collect self-reported data on feedback skills with six items (e.g., "I do not have difficulties to assess others' performance.") and on tutoring skills with fifteen items (e.g., "I first

make sure that the other person understands the assignment.").

Procedure

This study took place in the course of three Tuesdays spread over a period of three weeks. On the first Tuesday, participants received an introduction and took pre-measures: a prior knowledge test and two Likert scales on feedback skills and tutoring skills. On the second Tuesday, students who were to become tutors received a 70-minute training given by the two prime investigators of this study. Training was given to enhance either tutors' content knowledge or their tutoring skills. Immediately following the training, CK and TS tutors received their anonymized tutee's research questions together with the instruction on how to "Formulate feedback to help your tutee to improve his/her research questions." Tutors had 30 minutes to write feedback, being allowed to consult the training materials. On the third Tuesday, tutees received their anonymized tutor's feedback. Tutees first had to take 20 minutes to understand and interpret their tutor feedback by answering three open questions regarding difficult and uncertain feedback aspects, as well as aspects that needed further feedback. After answering these three questions, they had 40 minutes to revise the draft research question based on their tutor feedback. Finally, tutees had five minutes to fill in the Likert-scale about their motivation to revise research questions.

Scoring

Tutor performance on giving feedback

We selected a random subset of 14 out of 47 tutor feedback items, which was independently scored by two researchers (the first author and a colleague who was unaware of the treatments of this study). A three-point scoring rubric was used to assist in a reliable assessment of

performance, based on a selection of five general criteria for effective feedback used in other feedback studies (Narciss, 2008; Nelson & Schunn, 2009; Strijbos et al., 2010): coverage, identifications of problems, explanations for the problems identified, suggestions, and affective language (see Appendix). The maximum total score for tutor feedback performance was 15. The Pearson correlation coefficient was calculated for the interrater reliability on the scores of each criterion and the total scores: coverage, $r = 1.00$, $p < .01$; identifications of problems, $r = .55$, $p = .04$; explanations for the problems identified, $r = .64$, $p = .01$; suggestions, $r = .67$, $p = .01$; affective language, $r = .97$, $p < .01$; and the total scores, $r = .90$, $p < .01$. For the purpose of this research, these correlations indicated a satisfactory level of consistence between two researchers in ranking tutor feedback based on each criterion and total scores. One researcher scored all of the 47 tutor feedback and those scores were used for the analyses.

Tutee performance on research questions

We selected a random subset of 14 out of 47 research questions, which was independently scored by two researchers (the third author and a colleague) based on an overall impression of the quality. Both of them were unaware of the treatments of this study. The maximum score for each research question was 20. Pearson correlation coefficient was calculated for the interrater reliability of the research question scores, $r = .73$, $p < .01$. For the purpose of this research, this indicated a satisfactory level of consistence between two researchers in ranking the quality of tutee research questions. One researcher scored all of the 47 research questions and these scores were used for the analyses.

Analyses

Different statistical tests were applied to analyze each dependent variable depending on the type

of data, design, and whether assumptions for the test could be met. A significance level of .05 was used for all analyses.

Results

Tutor feedback performance

Tests of normality revealed that tutor feedback score on each criterion and the total score were not normally distributed for both groups. Thus, the Mann-Whitney tests were used to compare mean ranks of CK and TS tutors' scores on the five criteria of effective feedback and the total scores. Table 2 shows the medians and ranges of CK and TS tutors' feedback scores. The test result on the total scores of feedback indicated that TS tutors gave more effective feedback than CK tutors. The effect size (r) estimate indicates that tutor training had a large (> .50) effect on the overall effectiveness of tutor feedback. As for test results on individual criteria, TS tutors provided more appropriate explanations to the problems identified, gave more suggestions to improve the problems identified and used more affective language than CK tutors. No significant differences were found in what the feedback covered and the number of problems identified between CK and TS tutors.

Tutee interpretations of tutor feedback

Regarding tutee interpretations, a substantial number of tutees did *not* answer the three interpretation questions properly. They either left the answers to these questions blank or simply answered "no": 36% of tutees did not indicate any difficult aspects of tutor feedback, 36% of tutees were not uncertain of any tutor feedback, and 23% of tutees did not need any further tutor feedback. The number of feedback aspects that tutees found difficult, uncertain and needed for

further feedback are quantified and summarized in Table 3. The values ranged from zero (no aspect) to two (two or more than two aspects). On average, both groups of tutees reported less than one aspect that they found difficult, uncertain or needed for further feedback.

Tutee performance on revised research questions

Table 4 shows means and standard deviations for the scores on the draft and revised research questions. In general tutees' performance on the draft and revised research questions was *not* satisfactory. There seems to be no substantial difference between CK and TS tutees in the scores of the draft and revision. An ANCOVA was applied to compare the revision means between CK and TS tutees by using the draft scores as a covariate in order to control for their influence on the revision scores. The results showed that the covariate, draft score, was significantly related to the revision score, $F(1,44) = 7.77$, $p = .01$. However, no significant difference was found in the revision scores between CK and TS tutees after controlling for the effect of draft score, $F(1,44) = .003$, ns. (Note that while values of the F statistic less than 1 can occur by chance when the null hypothesis is true or near true ~~as others have explained~~, values close to 0 can indicate violations of the assumptions that ANCOVA depends on. The assumptions of test of normality, independence of the covariance and treatment effect, homogeneity of regression slopes, and Levene's test of equality of group variances were examined prior to the ANCOVA test. But the test results showed that none of these assumptions was violated.)

Tutee motivation to revise

As to tutee motivation to revise, the *Cronbach's* alpha coefficient for the eight items is .90, suggesting that the items had a satisfactory internal consistency regarding tutees' attitudinal data on motivation to use tutor feedback to revise their research questions. Since a Likert scale was

used to measure tutee motivation to revise, an independent-samples *t*-test was conducted to compare the overall means of the eight items between CK ~~(*M* = 3.27, *SD* = .86)~~ and TS ~~(*M* = 3.81, *SD* = .71)~~ tutees (Table 5). There was a significant difference in the overall means, $t(43) = -2.30$, $p = .03$, $r = .33$, suggesting that tutor feedback had an medium (> .30) effect on tutee motivation to revise. Specifically, results suggest that tutees who received feedback from TS tutors were more motivated to revise their research questions than those who received feedback from CK tutors.

Pre-measures

Table 6 shows the means and standard deviations for CK and TS groups' pre-measures. To detect a possible lack of similarity between CK and TS groups, two separate MANOVA tests were conducted for the tutees and tutors' scores on the three pre-measures of content knowledge, feedback skills and tutoring skills. Based on the tests results, no significant differences were found, neither between CK and TS tutees, $T = .02$, $F(3,43) = .34$, $p = .79$, $\eta_p^2 = .02$, nor between CK and TS tutors, $T = .01$, $F(3,43) = .08$, $p = .97$, $\eta_p^2 = .01$. Univariate tests revealed that none of these pre-measures was significantly different.

Discussion and Conclusions

The tutoring task of this study, formulating feedback on tutee research questions, required peer tutors to apply both content knowledge ~~in~~to the domain of formulating research questions and tutoring skills of giving feedback. Our results not only showed that TS tutors formulated more effective feedback than CK tutors but also that TS tutees had a higher level of motivation to revise than CK tutees. These findings support the claim that pedagogical knowledge in tutoring skills is a more critical component of effective tutoring on complex tasks than content knowledge

in a particular domain (Falchikov, 2001; Lepper & Woolverton, 2002). In addition, based on the training materials, one might expect that either group of tutors would perform better than the other on some of the five general criteria of effective feedback. Surprisingly, CK tutors did not perform significantly better in covering the main and sub-questions, identifying the problems in their tutees' research questions and giving explanations, even though specific attention was drawn to what makes good research questions during the training. In contrast, TS tutors not only performed significantly better in giving suggestions for improvement and using affective language but also on giving explanations even though they were only trained in how to give feedback. One may speculate that the pedagogical knowledge of tutoring skills also triggers any tutors' existing knowledge when performing tutoring tasks.

Still, we did not find any evidence for the effect of tutor feedback on tutee performance of revising their research questions. Our results on tutor and tutee data seem to be inconsistent: if the training did influence tutor feedback performance, and resulted in different levels of tutee motivation to revise, why was there then no difference in tutee revision performance?
Other tutoring studies which trained students in advance (Fuchs et al., 1994; King, 1991, 1994; King et al., 1998; Nath & Ross, 2001) show, as do our findings that training influences how well peer tutors perform their tutoring tasks. However, the absence of significant effects of tutor feedback on tutee revision performance seems to deviate from the performance results found in King's studies (1991, 1994; 1998). Our deviating results may be explained by a different set-up designed to prevent reciprocity effects, in which students took on a fixed role of being either tutor or tutee to perform their assigned role task. This set-up aimed to rule out practice effects and the additional cognitive benefits caused by reciprocal tutoring. Corresponding to this non-reciprocal set-up, *only* tutors were trained with the treatment and tutees did *not* receive any

training at all. However, as mentioned, the effectiveness of tutor feedback not only depends on how tutors give feedback but also depends on whether tutees attend to the feedback and how they process the feedback. Without instructional interventions on a mindful reception or strategies of processing, tutees might not attend to or not act upon tutor feedback (Gielen, Peeters, Dochy, Onghena, & Struyven, 2010; Narciss, 2008). Though we did ask tutees to first interpret tutor feedback prior to revising their research questions, tutees' limited responses to the three interpretation questions indicate that they did not properly attend to tutor feedback. This might explain why there was no effect of tutor feedback on tutee revision performance.

Another possible explanation might be the difficulty for tutees to understand and use written feedback in their revision. As indicated by studies on the effectiveness of written feedback, students find written feedback difficult to understand and they cannot use the information to fill the gap between their performance and tutor judgments (Blair & McGinty, 2012; Van der Schaaf, Baartman, Prins, Oosterbaan, & Schaap, 2011). In particular, tutees in our study were not able to ask their tutors for clarifications prior to revision and in turn tutees might not understand, use or act upon tutor feedback in their revision.

There are three other issues that also need to be explored. The first issue concerns generalizability. Without a control group, the findings of this study can only be generalized to a similar set-up with two groups (i.e., tutors with content knowledge or tutoring skills). To extend the generalizability of this study, it would be desirable to compare the effects of content knowledge or tutoring skills to a control (without any training) or a combination (both content knowledge and tutoring skills) group. The second issue concerns validity. The independent variable was tutor training in content knowledge or tutoring skills. Validity could be increased by ensuring that tutors indeed had content knowledge in a particular domain or pedagogical

knowledge of tutoring skills by measuring their attainments of the training. The third issue concerns the scope of the training effects. In this study, peer tutors were only asked to formulate written feedback after the training. Further research should also examine how two types of training influence tutors in conducting an interactive feedback dialogue with their tutees to help them revise research questions. It would be interesting to find out whether tutors trained in content knowledge result in different types of dialogues from those trained in tutoring skills. In addition, it is very likely that tutors who received instructions on formulating research questions would apply the training content to improve their own research questions. It would be valuable to examine such training effects on tutor learning in future research.

References

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives*. New York, NY: Longman.

Blair, A., & McGinty, S. (2012). Feedback-dialogues: Exploring the student perspective. *Assessment & Evaluation in Higher Education, 38*(4), 466-476. doi: 10.1080/02602938.2011.649244

Chi, M. T. H., Siler, S. A., & Jeong, H. (2004). Can tutors monitor students' understanding accurately? *Cognition and Instruction, 22*(3), 363-387. doi: 10.1207/s1532690xci2203_4

Davis, W. K., Nairn, R., Paine, M. E., Anderson, R. M., & Oh, M. S. (1992). Effects of expert and non-expert facilitators on the small-group process and on student performance. *Academic Medicine, 67*(7), 470-474.

De Grave, W. S., De Volder, M. L., Gijselaers, W. H., & Damoiseaux, V. (1990). Peer teaching and problem-based learning: Tutor characteristics, tutor functioning, group functioning and student achievement. In Z. M. Nooman, H. G. Schmidt & E. S. Ezzat (Eds.), *Innovation in medical education: An evaluation of its present status* (pp. 123-134). New York, NY: Springer.

Falchikov, N. (1995). Improving feedback to and from students. In P. Knight (Ed.), *Assessment for learning in higher education* (pp. 157-166). London, England: Kogan Page.

Falchikov, N. (2001). *Learning together: Peer tutoring in higher education*. New York, NY: Routledge.

Fuchs, L. S., Fuchs, D., Bentz, J., Phillips, N. B., & Hamlett, C. L. (1994). The nature of student interactions during peer tutoring with and without prior training and experience. *American Educational Research Journal, 31*(1), 75-103.

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. doi: 10.1016/j.learninstruc.2009.08.007

Jenkins, J. R., & Jenkins, L. M. (1987). Making peer tutoring work. *Educational Leadership, 44*, 64-68.

King, A. (1991). Effects of training in strategic questioning on children's problem-solving performance. *Journal of Educational Psychology, 83*(3), 307-317.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368. doi: 10.3102/00028312031002338

King, A. (1997). ASK to THINK-TEL WHY: A model of transactive peer tutoring for scaffolding higher level complex learning. *Educational Psychologist, 32*(4), 221-235.

King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Review, 10*(1), 57-74.

King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology, 90*(1), 134-152.

Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the study of highly effective tutors. In J. M. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 135-158). New York, NY: Emerald Group Publishing.

Moust, J. H. C., & Schmidt, H. G. (1994). Effects of staff and student tutors on student achievement. *Higher Education, 28*(4), 471-482.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. Van Merriënboer & D. M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 125-144). New York, NY: Routledge.

Narciss, S., & Huth, K. (2006). Fostering achievement and motivation with bug-related tutoring feedback in a computer-based training for written subtraction. *Learning and Instruction, 16*(4), 310-322. doi: http://dx.doi.org/10.1016/j.learninstruc.2006.07.003

Nath, L. R., & Ross, S. M. (2001). The influence of a peer-tutoring training model for implementing cooperative groupings with elementary students. *Educational Technology Research and Development, 49*(2), 41-56.

Nelson, M., & Schunn, C. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science, 37*(4), 375-401. doi: 10.1007/s11251-008-9053-x

Schmidt, H. G., Van der Arend, A., Moust, J. H. C., Kokx, I., & Boon, L. (1993). Influence of tutors' subject-matter expertise on student effort and achievement in problem-based learning. *Academic Medicine, 68*(10), 784-791.

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. doi: 10.1016/j.learninstruc.2009.08.008

Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education, 32*(3), 321-345.

Van der Schaaf, M., Baartman, L., Prins, F., Oosterbaan, A., & Schaap, H. (2011). Feedback dialogues that stimulate students' reflective thinking. *Scandinavian Journal of Educational Research, 57*(3), 227-245. doi: 10.1080/00313831.2011.628693

Van der Schee, J. (2001). How to train students to formulate good research questions? *International Research in Geographical and Environmental Education, 10*(3), 245-259. doi: 10.1080/10382040108667444

Appendix: Scoring rubrics for tutor feedback performance

	1	2	3
Coverage Whether the feedback referred to the main- and sub-questions	referring to only the main or to only the sub-question(s)	referring to the main and one of the sub-questions	referring to both main and two sub-questions
Identification of the problems Whether problems of tutees' research questions were properly identified For example, a question formulated with a yes/no question is identified as a problem of being NOT FEASIBLE. Then this problem is identified wrongly. Because a yes/no question is not related to whether a question is feasible or not.	only one or two problems were identified or they were identified incorrectly	half of the problems were identified correctly	two-thirds of the problems were identified correctly
Explanations for the problems identified Whether appropriate explanations were given to the problems identified	no explanation were given for the problems identified or none of the given explanations matched the problems identified	explanations were given to half of the problems identified and half of the given explanations matched the problems identified	explanations were given to two-thirds of the problems identified and half of the given explanations matched the problems identified
Suggestions Whether suggestions for improvements were given to the problems	no suggestion was given for the problems identified	suggestions were given to half of the problems identified	suggestions were given to two-thirds of the problems identified
Affective language Whether two types of affective language were used in the feedback: i) praise, summary or conclusion used to state what has been done well; and ii) mitigating language used in *identification of the problems* or in *suggestions*	no affective language is used simply describing that a research question meets or does not meet a certain criterion	only one type of affective language is used	both types of affective language are used

Table 1. An example of tutee draft research question

Theme	Biosphere, sea level rise
Main-question	Would the temperature drop by the melting of the ice caps, because a shorter trade route is created which reduces CO2 emissions?
Sub-question 1	How long does it take for the caps to melt enough so that a new trade route through the poles can emerge?
Sub-question 2	Is the trade route so much shorter and faster so that the pollution on earth drops?

Table 2. Medians (Range)and Mann-Whitney tests for tutor feedback performance

	Coverage	Problems	Explanations	Suggestions	Affective language	Total
CK	3.00 (2)	2.00 (3)	2.00 (3)	1.00 (3)	1.00 (2)	9.00 (10)
TS	3.00 (1)	1.00 (3)	3.00 (1)	2.00 (2)	3.00 (2)	12.00 (8)
U	218.00	236.00	131.00	92.50	51.00	72.00
z	-1.86	-.90	-3.42	-4.16	-5.36	-4.37
p	.07	.40	< .001	< .001	< .001	< .001
r	-.27	-.13	-.50	-.61	-.78	-.64

Table 3. Means (Standard deviations) for the number of aspects that tutees found difficult, uncertain and needed for further feedback

	CK	TS
difficult aspects to understand (0-2)	.73 (.63)	.56 (.65)
uncertain aspects (0-2)	.86 (.77)	.76 (.66)
aspects needed for further feedback (0-2)	.68 (.84)	.80 (.71)

Table 4. Means and Standard Deviations for tutee performance on the draft and revision

	All tutees		CK tutees		TS tutees	
	M	*SD*	*M*	*SD*	*M*	*SD*
draft (1-20)	9.38	2.88	9.27	3.09	9.48	2.74
revision (1-20)	9.38	2.82	9.32	2.55	9.44	3.10

Table 5. Means and Standard Deviations for tutee evaluation of tutor feedback

Item	treatment	*M*	*SD*
I am satisfied with the feedback.	CK	3.05	1.25
	TS	3.74	.96
The feedback is written clearly and easy to understand.	CK	3.41	1.05
	TS	3.96	.88
The feedback aims to improve my research questions.	CK	3.45	1.14
	TS	4.00	.74
The feedback encourages me to improve my research questions.	CK	2.95	1.05
	TS	3.57	1.04
My tutor gives me positive feedback.	CK	3.09	1.31
	TS	4.09	1.00
My tutor has explained clearly what was not good in my research questions.	CK	3.23	1.19
	TS	3.52	.99
I accept the feedback pointing to what was not good in my research questions.	CK	3.77	1.11
	TS	3.96	.88
I find the suggestions of my tutor useful to improve my research questions.	CK	3.23	1.07
	TS	3.65	.94
overall means	CK	3.27	.86
	TS	3.81	.71

Table 6. Means (Standard Deviations) and MANOVA tests for pre-measures

	Tutees			Tutors		
	prior knowledge test (tot: 10)	feedback skills (tot: 5)	tutoring skills (tot: 5)	prior knowledge test (tot: 10)	feedback skills (tot: 5)	tutoring skills (tot: 5)
CK	4.45 (1.61)	3.64 (.66)	3.64 (.56)	4.18 (1.13)	3.70 (.47)	3.55 (.37)
TS	4.32 (1.42)	3.68 (.44)	3.57 (.33)	4.08 (1.22)	3.73 (.71)	3.73 (.71)
F (1,45)	.09	.07	.35	.18	.03	.03
p value	.76	.79	.56	.67	.87	.87
η_p^2	0	0	.01	0	0	0