De Wever, B., Hämäläinen, R., Voet, M., & Gielen, M. (2015). A wiki task for first-year university students: The effect of scripting students' collaboration. The Internet and Higher Education, 25, 37–44. doi:10.1016/j.iheduc.2014.12.002

Authors

Bram De Wever: http://www.tecolab.ugent.be/pages/bram.html

Raija Hämäläinen: http://www.tecolab.ugent.be/pages/raija.html

Michiel Voet: http://www.tecolab.ugent.be/pages/michiel.html

Mario Gielen: http://www.tecolab.ugent.be/pages/mario.html

A wiki task for first-year university students: the effect of scripting students' collaboration

Bram De Wever, Raija Hämäläinen, Michiel Voet, and Mario Gielen

Abstract

Developing authentic learning environments in higher education calls for pedagogical approaches to foster online collaborative learning. The main aim of this study was to investigate the effect of a collaboration script for a wiki task. A collaboration script is a set of instructions to improve collaboration between learning partners. Participants were first-year university students in Educational Sciences (N=186) collaborating in groups of five during a three-week period to create a wiki on peer assessment in education. Two conditions were contrasted: a scripted and a non-scripted condition. The effect of scripting was measured in four ways (questionnaires, log-file analyses, group product scores, and individual pre- post-test scores). Results show significant positive effects of scripting with respect to the collaborative group processes and students' feelings of shared responsibility. No significant effects of scripting were found with respect to the developed wiki products. As for students' individual learning outcomes, results showed a significant increase from pre- to post-test for all students. Although the increase was higher in the scripted condition, the difference between the conditions was not statistically significant.

Keywords

Wiki, script, collaboration, collaborative learning

A wiki task for first-year university students: the effect of scripting students' collaboration

1. Introduction

1.1. Wikis in higher education

Web 2.0 applications are often suggested as great tools for shaping educational practices (for an overview, see Hsu, Ching, & Grabowski, 2014). These applications appear to be especially useful to prepare students for future work environments, which are characterized by an increasing demand for advanced skills to analyze information and solve complex problems in inter-professional groups (Noroozi, Biemans, Weinberger, Mulder, & Chizari, 2013; Tynjälä, 2013). More specifically, the implementation of social software and collaborative learning methods allows the creation of learning environments in which authentic tasks resembling those in professional contexts are simulated (Tynjälä, Häkkinen, & Hämäläinen, 2014).

Recently, the educational potential of wiki-environments has been widely discussed (see e.g. De Wever, Van Keer, Schellens, & Valcke, 2011; Ertmer, Newby, Yu, Liu, Tomory, et al., 2011; Xiao & Lucking, 2008; West & West, 2009). At the general level, the advantages of wikis are typically described as being helpful in organizing learning activities (e.g. for setting up collaborative learning spaces) and having a positive influence on learning (outcomes or shared processes). In practice, wikis are seen as effective tools for collaborating on shared documents (Kear, Woodthorpe, Robertson, & Hutchison, 2010), for example, in the context of international collaboration (Ertmer, Newby, Liu, Tomory, Yu, 2011). In addition, wiki environments have the potential to support students in developing new skills in conjunction with their peers (Lai & Ng, 2011). In particular, Laru, Näykki, and Järvelä (2012) found that shared use of the wiki-environments to perform multiple tasks might improve individual knowledge acquisition.

Despite the advantages mentioned, the educational use of wikis also raises some critical questions. The other side of the coin is that, although wikis may be useful tools for higher education, there seems to be a vast dissimilarity in students' experiences. While some students highlight wikis' ability to support learning, by helping to organize and coordinate thoughts, other students report problems resulting from group collaboration (e.g. difficulties in engaging in shared group processes) and frustrations related to the wiki software used (Meyer, 2010). In a study by Wheeler, Yeomans, and Wheeler (2008), first-, second- and third-year bachelor students in an undergraduate teacher training program were asked to post their views on the use of wikis. Interestingly, while students were positive about the idea of sharing their writings with others, they did not like the idea that their fellow students could edit their contributions. More specifically, such reservations were especially demonstrated by first-year students. Similarly, Kale (2013) found that some learners feel uncomfortable editing others' ideas in wikis. Students, especially first-year students in higher education, may thus be rather reluctant to change each other's ideas and thoughts or comment upon them (De Wever, 2011).

With regard to learning in wiki-environments, triggering productive collaborative learning may be challenging, as several studies have reported problems concerning shared learning processes, caused by unequal participation (e.g. O'Bannon, Lubke, & Britt, 2013). For example, Wheeler et al. (2008) claim that "students tend to read only those pages to which they had contributed, which tend to negate the original objective of collaborative learning through content generation" (p. 993). This means that students may not engage in productive activity in terms of co-constructing knowledge, by building on each other's work and editing each other's pages. Therefore, one of the major challenges in applying wikis in an educational setting seems to be how to trigger and maintain productive group processes. This is a particularly challenging task, as wiki's are often implemented for distributed learning

activities, taking place in an online environment. Previous research shows that students do not spontaneously form online learning communities, but often experience feelings of isolation and disconnection among each other (Boling, Hough, Krinsky, Saleem, & Stevens, 2012).

1.2. Rationale for this study and research questions

One of the main questions driving this research is how to arrange learning activities and how to foster collaboration in a wiki-environment. More specifically, the rationale for this study can be found in earlier findings (De Wever, 2011), indicating that students may be reluctant to change or comment upon each other's work. Instead of working collaboratively and constructing knowledge based on each other's ideas and thoughts in the wiki-environment, students are often working independently on specific pages, i.e. they are each focusing on a subtask.

Previous findings have indicated that instructional support is needed in order to generate high-level collaborative activity and to acquire adequate collaboration skills (Cole, 2009). In this respect, collaboration scripts have been introduced as a way to bring about productive group processes and shared work, in which learning situations are prearranged and designed in a way that shared problem solving is triggered naturally (for detailed description see, Kobbe, Weinberger, Dillenbourg, Harrer, Hämäläinen, et. al., 2007). A collaboration script is a set of instructions to improve collaboration through structuring the interactive processes between learning partners, by organizing the task and the collaborative process (Kollar, Fischer, & Hesse, 2006). Several studies have reported the positive effects of such scripts (see e.g. the review study by Fischer, Kollar, Stegmann, & Wecker, 2013). Despite the potential of scripts, so far only a limited number of the studies have focused on applying collaboration scripts in wiki environments. Recently, Wichmann and Rummel (2013) argued

that a scripting approach has the potential for fostering collaboration during wiki-based writing.

The main aim of this study is to explore the implementation of a wiki task in higher education, and more specifically to study the effect of applying a collaboration script for this wiki task. The central goal of collaboration scripts is to shape the way learners build knowledge based on each other's ideas and thoughts (see, Kobbe, et. al., 2007). In practice, scripts operate by sequencing activities and assigning roles towards effective collaborative processes, related to a thoughtful use of the available resources and/or task division. The main goal of the script introduced here, is to enhance students' collaboration and to increase their feelings of shared responsibility for the full task and, as such, to enhance students' collaborative learning. In our study, two conditions (scripted versus non-scripted groups) were implemented. The workings of the script and the differences between conditions will be explained in detail in the method section.

The following research questions and corresponding hypotheses are put forward:

RQ1: Is there a difference between students' experiences regarding the collaboration in the scripted versus non-scripted groups? It was expected that the script would increase students' reviewing and editing of each other's work. Therefore the following specific hypotheses were formulated: Students in the scripted condition report that they (1) read more of the pages in the wiki, (2) edited more of the pages in the wiki, (3) tackled the work more together, (4) felt themselves more responsible for the complete wiki, and that (5) were more inclined to feel the whole group was responsible for the wiki.

RQ2: Is there a difference between students' behavior in the scripted versus non-scripted groups? More specifically, can we observe differences in the log files of the wiki that can confirm students' self-reported experiences (see RQ 1)? Hypotheses here are: (1) students

in the scripted conditions are editing more pages in the wiki, and (2) are taking more turns editing the pages.

RQ3: Is there a difference between the two conditions with respect to the quality of the product (i.e. the developed wiki)? Given that some literature sees merit in task specialization, whereas other research sees more benefit in shared collaborative processes, no specific hypotheses are formulated here.

RQ4a: Does the collaborative work lead to an increased (content) knowledge in all (i.e. scripted and non-scripted) groups? As Laru et al. (2012) argued that shared use of wikis may improve individual knowledge acquisition, and given that students were actively reading sources and processing information on their wiki pages, we hypothesize that students' content knowledge will be increased after the wiki task.

RQ4b: Is there a differential impact of scripting on students' content knowledge? In other words: is there a difference between the content knowledge of students in the scripted groups and students in the non-scripted groups? Since scripts have been shown to be beneficial for collaborative learning processes (Fischer et al., 2013), we hypothesize that students' content knowledge in scripted groups will be higher.

2. Material and methods

2.1. Context and participants

The participants in this study were first-year university students Educational Sciences (N= 186) taking the course Instructional Sciences. Participation in the wiki-assignment was a complimentary part of the course. Students were randomly assigned to a group. Due to non-participation (drop-out) 4 groups consisted of 4 students, while the other 34 groups consisted of 5 students. In total, 186 students were divided over 38 groups, of which 1 student did not fill in the pre-test and 10 students did not fill out the post-test and post-questionnaire.

2.2. Task

Students had to work together during a three-week period in order to create a wiki documenting the use of peer assessment in education. The same case was presented to each group of students: “A number of lecturers at a University College want to implement peer assessment in their courses. During a team meeting, they realize that a lot of colleagues still have some questions concerning this particular form of evaluation. In order to prepare its implementation, you and your group members are asked to jointly prepare a wiki documenting this assessment practice. Specific questions that should be addressed are: What is peer assessment exactly? Why should it be used? Does it work well? Is it an “honest” form of assessment? What has research shown? How is this form of assessment related to other forms? When is peer assessment effective? Does it work for all students? What about the time investment of lecturers and students? How do students experience peer assessment?”

All groups were required to develop an informative wiki containing the following pages on peer assessment: a page for the (1) overview, (2) description, (3) theoretical rationale, (4) advantages, (5) disadvantages, and (6) points of attention when implementing peer assessment in educational practice. Students were provided with 10 external sources. These were 10 research articles, of which 5 were labeled as “main sources” and 5 as “additional sources”. In addition, they were informed that seeking and using additional resources was allowed.

2.3. Research design

Two conditions were contrasted within this study: a scripted and a non-scripted condition. Groups were randomly assigned to one of the conditions. In the non-scripted condition, students were asked to study the provided resources, the main sources being the most

important ones, and to develop the required wiki pages. In this condition, students were free to organize their group work. In contrast to this, students in the scripted condition were suggested to follow a script. Specific guidelines, in the form of a step-by-step plan, were provided to organize their group work in order to ensure that students were actually building new knowledge on each other's work (Arvaja, Salovaara, Häkkinen, & Järvelä, 2007). Thus the aim of the script was to foster students to engage in productive activity in terms of co-constructing knowledge. Students were informed that these guidelines were meant to guide them and that they were to be seen as a flexible aid and not as a strict path (cf. macro-scripts as pedagogical models to enhance group work by providing group members with general guidelines to organize and sequence productive collaborative activities (e.g. inquiry cycle), Dillenbourg & Tchounikine, 2007). The first step of the script required each of the students to read two different sources (one of the main and one of the additional sources) and each student was suggested to start drafting a different page of the required wiki pages. In this respect, the external resources as well as the pages to write a draft for, were divided among students at the start of the task, in such way that all students initially started reading different sources and started writing drafts for different wiki-pages. After this, the second step required the first student of the group to read another one of the main sources (see Figure 1) and to edit the wiki page that was drafted by the second student in the group (see also Figure 1). At the same time, this second student had to read another one of the main sources and continue working on the page that was initially drafted by the third student, by constructing the existing draft further. Similarly, the third, fourth, and fifth student had to read another main source, and continue working on the page initially drafted by respectively the fourth, fifth, and first student. In step 3, the same procedure was repeated, but now the first student had to read yet another source and further construct the page the third student initially drafted, etc., in such way that every student was now working on another draft. This procedure continued through

step 4 and 5, but again students had to read another source and contribute to another draft. In this way the script was expected to stimulate students to edit others' ideas in wikis and thus to increase the amount of work shared. At the end, students had to finalize the wiki-page they drafted in the first step. Students had about three to four days for each step.

Eventually, students in the group had at least read the main sources, edited all the required wiki-pages (mentioned under 2.2. above), and were responsible for finalizing the page they originally drafted during step 1. A detailed overview of this script is presented in Figure 1. Please note that although the script required students to co-construct knowledge on each of the wiki pages, in line with the idea of fostering shared responsibility, it did not require students to read all of the additional sources.

<< Figure 1. Overview of the six steps suggested in the collaboration script >>

2.4. Data collection and measures

For research question 1, students were asked to fill in a questionnaire after the intervention. This questionnaire consisted of 7-points Likert scale items (from completely disagree to completely agree) investigating the extent to which students tackled the work together, read all wiki pages, felt responsible, and wrote together on a wiki-page (see Table 3 for further details). In addition, three self-report control variables were measured. We asked students (1) how they experienced the collaboration within their group (on a scale from 0 to 10), (2) how many hours they spent on the wiki assignment, and (3) how many of the given sources they read.

For research question 2, a second source of information was used. Log files of the system were used to analyze students' collaboration behavior. More specifically, the total

amount of edits, the number of separate pages edited, and the amount of turn-takes were calculated from the log files and further analyzed to answer this research question.

For research question 3, we used the wiki grades. The wikis were graded by the teaching assistant who was responsible for grading all wikis related to this course, and has been doing this for several years. This teaching assistant was aware that a research study was set up. However, although she had full access to the wiki and the wiki history, she was not aware of the condition the wikis belonged to at the moment of grading, in order to avoid a possible bias. Wikis were scored on four aspects (content, depth, structure, and style), and a combined score (scale 0-100) was calculated.

For research question 4, each student was asked to answer the same 5 questions individually during a pretest and a posttest. The post-test was administered together with the questionnaire (cf. RQ1). Students were informed that we were gauging their knowledge on peer assessment. Instructions for the pre-test included the information that it was perfectly possible (and no problem) that students were unable to answer all questions. Instructions for the post-test included a notification that the scores would only be used for research and would not influence the grading of the students.

2.5. Content analyses of the pre- and post-test

In order to analyze students' answers on the five questions, a coding scheme was developed based on the sources given to the students (see Table 1 for details). The coding scheme contained a set of criteria for each question, against which provided answers were compared. First, this coding scheme was discussed by the four coders (the first, third, and fourth author of this article and one graduate student), after which these raters independently coded all answers. Answers were blinded and the order was randomized, so that the coders had no idea which student gave the answer or whether it was an answer from the pre- or the post-test.

Initial reliability was good for more than half of the categories (Krippendorff's Alpha higher than .75 in 14 out of the 24 categories). However, for the other categories only a fair agreement was found (see Table 1), we decided to discuss all codes with disagreement. In most cases, differences were due to coder mistakes, such as typos, coder fatigue (clear elements that were not noticed), or misinterpretation. Whenever that was the case, the disagreement was resolved. However, when no coder mistakes were made, coders kept their initial code, in order to reflect the true agreement (and disagreement) between coders using the particular coding scheme. As can be seen in Table 1, some categories almost reached full agreement, while others showed some more variation between the coders. However, the final Krippendorff's alpha values were all higher than .78 (see Table 1), which can be considered good agreement beyond chance (Krippendorff, 1980). Based on the content analysis, scores were calculated taking the mean of all categories (scored between 0 and 1), and multiplying this proportion by 100 to get a percentage score for each question (see the notes of Table 1 for more details).

<< Table 1. Overview of the questions and the reliability of the coded categories within the answer of each question. >>

2.6. Statistical Analysis

Given the hierarchical nature of the data (i.e. students nested in groups), multilevel analyses were applied to control for between group variance, using MLwiN 2.27. This was done for the questionnaire data (in order to answer RQ 1), the log-file data (RQ 2), the wiki scores (RQ 3), and the pre- post-test scores on content knowledge (based on the content analysis described in section 2.5., RQ 4).

3. Results

3.1. Research question 1: students' experiences (questionnaire data)

When students were asked to rate how they experienced the collaborative work for this specific assignment, they rated it on average 7.7 on a scale from 0 to 10 (see Table 2, a0) and no significant differences were found between the scripted (M=7.4) and the non-scripted (M=8.0) condition (p=.115, see Table 2, a1). For this dependent variable, no significant between-group variance was found (see Table 2, a0).

<< Table 2. Model estimates for the two-level analysis of students' experienced quality of the collaboration, time invested, and number of sources read. >>

For time investment, significant (p=.024) between-group variance was found (see Table 2, b0). When asked how many hours they spent working on this task, there was also a significant difference (p=.017) of more than 4 hours between the scripted condition (M=15.7) and the non-scripted condition (M=11.6, see Table 2, b1).

Also for the number of sources read, significant (p<.001) between-group variance was found. Students read on average 7.8 of the 10 provided sources (see Table 2, c0), and students in the scripted condition (M=9.3) read on average about three sources more (p<.001) than the non-scripted condition (M=6.3, see Table 2, c1).

Further on, Table 3 gives an overview of the average scores on the Likert-items focusing on the collaboration and responsibility. Differences were found between the conditions with respect to work division (both groups indicate that they divided the work, but scores for the non-scripted group were significantly higher) and tackling the work together (also here both groups indicate that they tackled the work together, however here the scripted group scored higher, see Table 3). In addition, students in the scripted group indicated more that they read all parts of the wiki, and that they felt responsible for all parts of the wiki. A

larger difference was found when asking students about who was responsible: students in the scripted condition agreed more with the statement that everybody was responsible for all parts of the assignment, whereas students in the non-scripted condition agreed more with the statement that each student was responsible for only one part of the wiki. Last but not least, the highest differences were found when asking students whether they helped with writing most parts of the wiki (scripted groups are agreeing, non-scripted groups are rather disagreeing) or wrote only on one part of the wiki (non-scripted group rather agreeing, scripted group rather disagreeing). For details on the averages and significance levels we refer to Table 3.

3.2. Research question 2: students' behavior (log-file data)

Details on the analyses of the log-file data are presented in table 4. More specifically, we analyzed the number of edits, the number of main pages edited, and the number of turn-takes on these five main pages. Every edit made on a wiki page by a different student was defined as a turn-take. Since we know from the results on the previous research question (as discussed in section 3.1.) that there is a difference in time investment between the scripted and the non-scripted conditions, we added time invested, together with the interaction effect of time invested with condition, as predictors in addition to the predictor condition. In this way, we control for the differences in time as found in the previous section.

<< Table 4. Multilevel analyses results of the log-file data. >>

When analyzing the total number of edits in the log files (see Table 4, a), no significant differences (p=1) with regard to the total number of edits between students in the scripted condition (M=51) and the non-scripted condition (M=51) were found. There was an

effect of time invested in the scripted (p=.023) and non-scripted (p=.005) condition, i.e. for an hour increase in time invested, students edited respectively 0.8 and 0.9 pages more on average. In this respect, there was no significant difference between both conditions (p=.796).

When we take a detailed look on the five pages of interest (i.e. the five main pages required in the task), we see this difference between the two conditions confirmed. Students in the scripted condition worked on average on 4.7 of the 5 pages, whereas in the non-scripted condition this is on average 3 (Table 4, b). Time invested is not significant in this model (p=.056 for non-scripted, p=.704 for scripted, difference between both conditions p=.321).

In addition to collaborating on more pages, students are also taking significantly (p<.001) more turns when working in scripted groups (M=7.4) than in non-scripted groups (M=4, cf. Table 4, c). Students in the scripted condition took more turns when they reported to have invested more time. This effect is small (for each hour more, about 0.1 turn-take) but significant (p=.010). This effect of time invested was not found for the non-scripted group (p=.899), leading to a significant (p=.040) difference between both conditions with respect to time invested (see interaction effect in Table 4, c).

3.3. Research question 3: the quality of the product (wiki scores data)

Research question 3 examines if there are differences between the final wiki product scores between the two conditions. Given that we know from the results of research question 1 that students in the scripted condition spent more time and read more sources, we controlled for these variables and their interaction effects with condition in our analyses. Results show no significant differences (p=.173; p=.993 after correction) between the total wiki scores of groups in the scripted (M=65.3; M=66.6 after correction and for a student with an average amount of time spent and an average amount of sources read) and the non-scripted (M=66.5;

M=66.7 respectively) condition. With regards to separate aspects (content, depth, structure, and style), no significant differences were found.

3.4. Research question 4: (differential) increase in students' content knowledge from pre- to post-test

Table 5 shows that in general, all students (overall, both conditions) increase their scores significantly from pre- to post-test. This also holds when we take a look at both conditions separately. When comparing the two conditions, we can notice that there are differences in increase from pre- to post-test between the two conditions, these are however non-significant for the total score and for three out of the five aspects. There were two aspects in which scripted students increased significantly more: the definition aspect and the added value aspect. For the latter, however, we did found significant differences between conditions at the start (lower scores for scripted condition).

<< Table 5. Mean estimates for pre-test, post-test, and increase for the total score and for each of the five questions separately, based on three level models (group, student, measurement occasion). >>

In order to investigate the differential impact of our script on students' post-test results, we modeled the post-test results, and checked the impact of the script. We controlled for pre-test scores, and again (given the differences found in RQ1) for time invested and sources read (see Table 6). As shown in Table 6, no significant differences are found for the total score. When looking at the separate aspects, only for the first question (definition), a significant difference between the two conditions is discovered, in favor of the scripted condition. Furthermore, we can notice an interaction effect of time invested and condition for

the background aspect: it seems that spending more time on the wiki is favorable for answering this question, but only for students in the scripted condition. Another interaction effect of sources read and condition for the added value aspect revealed that students in the non-scripted condition scored significantly higher when they actually had read more sources.

<< Table 6. Two level models of post-test results for the total score and each of the five questions separately (controlled for time invested and sources read). >>

4. Discussion

Despite the potential of social software, specific challenges may be encountered when new tools and methods are implemented in higher education practices. For example, Waycott, Sheard, Thompson, and Clerehan (2013) indicated that there is a tension between the participatory and collaborative nature of using social technologies, including wikis, on the one hand, and, on the other hand, the individual (and even competitive) nature of evaluating students. Another problem rising within groups in under-instructed environments is that courses often consist of groups that may reach high-level collaboration, and of groups in which shared work remains superficial or is entirely lacking (Hämäläinen & Häkkinen, 2010). Recently, Fischer et al. (2013) argued that learners with few successful collaborative experiences may not have adequate skills for productive collaboration in new learning contexts. Therefore, the development of higher education calls for pedagogical approaches, in which collaboration in social media settings is enabled and fostered. Specifically, there is the need for stimulating shared group processes because students do not spontaneously engage in editing others' ideas and joint collaboration in wikis (Kale, 2013). The main aim of this study was to explore the implementation of a wiki task in higher education, and more specifically to study the effect of applying a collaboration script for this particular wiki task. In general, the

results of this study coincide with the notions that wikis can enhance higher education. In both conditions (scripted and non-scripted), students' learning outcomes increased significantly from pre- to post-test.

Even though evidence for increasing learning outcomes can be derived from previous studies (Fischer et al., 2013), our results show no significant difference between the scripted and non-scripted groups' total wiki scores. Although individuals in the scripted groups scored higher on the post-test, this difference was not significant. Only for one of the five questions, students in the scripted groups scored significantly higher on the post-test. However, there was not a single case in which the increase between pre- and post-test was lower in the scripted condition than the control condition. Also, no significant differences were found regarding the final product of the group, i.e. the wiki they developed together. While our results did not show significant differences between the scripted groups' and non-scripted groups' total wiki scores and individual outcomes, the results based on our questionnaire data and log-file data did show interesting differences with respect to developing adequate group processes. Although students rated the quality of the collaboration the same in both conditions, students in the scripted condition stated that more time was invested and more sources were read compared to students in the non-scripted condition. Furthermore, the advantage of scripting is that students in the scripted groups felt more responsible for the complete wiki and indicated that they helped writing more pages than the non-scripted groups. This was confirmed by our log-file analyses; although no significant difference was found between the scripted and the non-scripted condition with respect to the total number of edits, students in the scripted groups worked on more pages and took more turns on these pages than the non-scripted students. From these findings, it can be concluded that the students in the non-scripted groups worked on fewer pages than those in the scripted groups. Finally, as the students in the scripted groups read more sources than the non-scripted

students throughout the task, altering the scripting approach may be one useful way to shape collaboration practices in wiki environments.

Kollar, Ufer, Reichersdorfer, Vogel, Fischer, and Reiss (2014) recently reported that while the collaboration script approach has been shown to be effective (Fischer et al, 2013), its effectiveness was specifically found in fostering social-discursive aspects of – in their case – argumentation skills, but scripts "rarely had additional positive effects on domain-specific outcomes" (p. 23). A similar reasoning may hold for our present study: the script is clearly beneficial for the collaborative processes (cf. RQ 1 and 2), but has no impact on the final product or content knowledge (cf. RQ 3 and 4).

In sum, based on our findings, there are two main reasons why scripting appears to be beneficial. First, it increases the shared responsibility within a collaborative environment. When responsibility for process and outcomes is shared more, it seems to also reflect on students' turn-taking behavior, i.e. taking turns on developing, reviewing, and rewriting of the same wiki-pages, instead of developing and rewriting single pages on their own. Second, the script can influence how thorough a task is dealt with, by raising the amount of external sources read and the amount of time spent on the task. In sum, our study showed that the scripting approach (Fischer et al., 2013) can be utilized as a starting point for developing practices that foster collaboration processes and joint problem solving in wiki environments (see also, Wichmann & Rummel, 2013).

4.1. Limitations, strengths, future research, and implications

This study was an attempt to investigate the influence of applying a collaboration script for a wiki task in higher education. The first limitation of our approach is that our study is a one-time event of three weeks. Therefore, further studies are needed to examine the potential of scripting for learning outcomes over longer time periods. A second limitation is that our

setting did not illustrate the variations and influences between the different learners within the scripted or non-scripted settings. For example, it seems plausible that, in some cases, the wikis may be fine-tuned by the most competent student(s) in the group, implying that the result of the collaborative writing process may heavily depend on the last specific edits. In addition to this, different mechanisms of collaborative writing (Lowry, Curtis, & Lowry, 2004; Onrubia & Engel, 2009) may be at play and current analyses reveal no detailed information on what students are exactly doing when they are taking turns. They could be editing each other's work, but also just adding ideas without changing the existing content, i.e. different collaborative writing strategies may be used and these may be an important factor explaining the results found. Therefore, future research could focus on in-depth analysis of the collaborative processes, by performing an interaction analysis of student activities in the online wiki environment, e.g. by focusing on argumentation (Kollar et al., 2014; Olson, Herbsleb, & Rueter, 1994) or on specific collaborative writing strategies used (Lowry et al., 2004; Onrubia & Engel, 2009). This type of detailed analysis of consecutive versions of several wiki pages could allow us to answer whether students are inserting additional text to qualify presented claims, adding alternative viewpoints, or focusing on deleting/inserting (in)correct content or claims. In this respect, future studies should explore more in detail which specific processes and text change operations (Southavilay, Yacef, & Calvo, 2010) are triggered, which collaborative writing strategies are used, and how they are influencing performance outcomes, taking into account the time used for the task. In addition to a more detailed view of the processes, these in-depth content analyses could provide a more fine-tuned measure for the quality of the end product.

Despite these limitations, our study has strengths. It sheds more light on two of the current challenges in the development of higher education. First, social media is creating new hopes for enhancing higher education. However, it is unclear how to bring about collaboration

in these new settings. Recently, Söderström, Häll, Nilsson, and Ahlqvist (2012) have argued that the quality of participants' activity rather than the new technologies brings about successful group activities. This study presents the scripting approach in wikis as one potential way to integrate research-based knowledge in authentic higher education practices and to stimulate collaboration. Second, a major research challenge in technology-enhanced learning is to find methodologically justifiable multiphase methods for developing a better understanding of collaboration. This study explores collaboration activities in wiki environments from several different perspectives, whereas previous research often focused on a single perspective. In more detail, our study investigates collaboration in wiki settings from the perspectives of students' experiences, students' behavior, students' content knowledge, and the quality of the product. Combining these methods (questionnaires, log-file analyses, group product scores, and individual pre- post-test scores) has rarely been done to date. While a lot of studies have been focusing on collaborative processes and the learning outcomes, few studies focused on students' experiences, feelings of responsibility, and related behavior, and even fewer studies combined all of these measures. In practice, this combination of methods has allowed us to gain a more in-depth understanding of collaboration in wikis with respect to developing higher education than one single method would.

Acknowledgments

The contribution of the second author was supported by the Academy of Finland (Project 258659).

The contribution of the fourth author was supported by Ghent University (BOF11/STA/026).

References

Arvaja, M., Salovaara, H., Häkkinen, P. & Järvelä, S. (2007). Combining individual and group-level perspectives for studying collaborative knowledge construction in context. *Learning and Instruction, 17*(4), 448-459.

O'Bannon, B, Lubke, J. & Britt, V. (2013). 'You still need that face-to-face communication': Drawing implications from preservice teachers' perceptions of wikis as a collaborative tool. *Technology, Pedagogy and Education*, *22*(2), 135–152.

Boling, E. C., Hough, M., Krinsky, H., Saleem, H., & Stevens, M. (2012). Cutting the distance in distance education: Perspectives on what promotes positive, online learning experiences. *Internet and Higher Education*, *15*(2), 118-126.

Cole, M. (2009). Using Wiki technology to support student engagement: Lessons from the trenches. *Computers and Education*, *52*(1), 141–146.

De Wever, B. (2011). Orchestrating collaborative learning in a wiki-environment. Paper presented at the 14th Biennial Conference for Research on Learning and Instruction (EARLI), Exeter, UK, 30 August - 3 September 2011.

De Wever, B., Van Keer, H., Schellens, T. & Valcke, M. (2011). Assessing collaboration in a wiki: The reliability of university students' peer assessment. *Internet and Higher Education, 14*(4), 201-206.

Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for computer-supported collaborative learning. *Journal of Computer Assisted Learning*, *23*(1),1-13

Ertmer, P.A., Newby, T.J., Yu, J.H., Liu, W., Tomory, A., Lee, Y.M., Sendurer, E., Sendurer, P., (2011). Facilitating students' global perspectives: collaborating with international partners using Web 2.0 technologies. *Internet and Higher Education, 14*(4), 251–261.

Ertmer, P.A., Newby, T.J., Liu, W., Tomory, A., Yu, H.E., & Lee, Y.M. (2011). Students' confidence and perceived value for participating in cross-cultural wiki-based collaborations. *Educational Technology, Research & Development, 59*, 213–228.

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.

Hämäläinen, R., & Häkkinen, P. (2010). Teachers' instructional planning for computer-supported collaborative learning: Macro-scripts as a pedagogical method to facilitate collaborative learning. *Teaching and Teacher Education, 26*(4), 871–877.

Hsu, Y., Ching, Y., & Grabowski, B. L. (2014). Web 2.0 Applications and Practices for Learning Through Collaboration. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology (pp. 747–758)*. New York, NY: Springer New York. doi:10.1007/978-1-4614-3185-5

Kale, U. (2013, iFirst). Can they plan to teach with Web 2.0? Future teachers' potential use of the emerging web. *Technology, Pedagogy and Education*. doi:10.1080/1475939X.2013.813408

Kear, K, Woodthorpe, J., Robertson, S. & Hutchinson, M. (2010). From forums to wikis: Perspectives on tools for collaboration. *Internet and Higher Education, 13*(4). 218-225.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., et al. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*(2/3), 211–224.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration Scripts – A Conceptual Analysis. *Educational Psychology Review, 18*(2), 159–185. doi:10.1007/s10648-006-9007-2

Kollar, I., Ufer, S., Reichersdorfer, E., Vogel, F., Fischer, F., & Reiss, K. (2014). Effects of collaboration scripts and heuristic worked examples on the acquisition of mathematical argumentation skills of teacher students with different levels of prior achievement. *Learning and Instruction, 32*, 22–36. doi:10.1016/j.learninstruc.2014.01.003

Krippendorff, K. (1980). *Content Analysis, an introduction to its methodology*. Thousand Oaks, CA: Sage Publications.

Lai, Y. C., & Ng, M. W. (2011). Using wikis to develop student teachers' learning, teaching, and assessment capabilities. *Internet and Higher Education, 14*(1), 15-26.

Laru, J., Näykki, P., & Järvelä, S. (2012). Supporting small-group learning using multiple Web 2.0 tools: A case study in the higher education context. *Internet and Higher Education*, *15*(1), 29–38.

Lowry, P. B., Curtis, A., & Lowry, M. R. (2004). Building a Taxonomy and Nomenclature of Collaborative Writing to Improve Interdisciplinary Research and Practice. *Journal of Business Communication, 41*(1), 66–99. doi:10.1177/0021943603259363

Meyer, K. (2010). A comparison of Web 2.0 tools in a doctoral course. *Internet and Higher Education, 13*(4), 226-232.

Noroozi, O., Biemans, H., Weinberger, A., Mulder, M. & Chizari M. (2013). Scripting for construction of a transactive memory system in multidisciplinary CSCL environments. *Learning and Instruction, 25*(1), 1–12.

Olson, G. M., Herbsleb, J. D., & Rueter, H. H. (1994). Characterizing the sequential structure of interactive behaviors through statistical and grammatical techniques. *Human-Computer Interaction*, 9(4), 427-472.

Onrubia, J., & Engel, A. (2009). Strategies for collaborative writing and phases of knowledge construction in CSCL environments. *Computers & Education, 53*(4), 1256–1265. doi:10.1016/j.compedu.2009.06.008

Southavilay, V., Yacef, K., & Calvo, R. A. (2010). Analysis of Collaborative Writing Processes Using Hidden Markov Models and Semantic Heuristics. In *2010 IEEE*

International Conference on Data Mining Workshops (pp. 543–548). IEEE. doi:10.1109/ICDMW.2010.118

Söderström, T., Häll, L., Nilsson, T., & Ahlqvist, J. (2012) "How does collaborative 3D screen-based computer simulation training influence diagnostic skills of radiographic images and peer communication?" *Contemporary Educational Technology*, *3*(4), 293-307.

Tynjälä, P. (2013). Toward a 3-P model of workplace learning: a literature review. Vocations and Learning. *Studies in Vocational and Professional Education*, *6*(1), 11–36.

Tynjälä, P., Häkkinen, P., & Hämäläinen, R. (2014). TEL@work: Toward integration of theory and practice. *British Journal of Educational Technology, 45*(6), 990–1000. doi:10.1111/bjet.12164

Waycott, J., Sheard, J., Thompson, C., & Clerehan, R., (2013). Making students' work visible on the social web: A blessing or a curse?, *Computers & Education, 68*, 86–95.

West, J. A., & West, M. L. (2009). *Using wikis for online collaboration: The power of the read-write Web*. San Francisco, CA: Jossey-Bass.

Wheeler, S., Yeomans, P., & Wheeler, D. (2008). The good, the bad and the wiki: Evaluating student-generated content for collaborative learning. *British Journal of Educational Technology, 39*(6), 987–995.

Wichmann, A., & Rummel, N. (2013). Improving revision in wiki-based writing: Coordination pays off. *Computers & Education, 62*, 262-270.

Xiao, Y. & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *Internet and Higher Education*, *11*(3-4), 186–193.

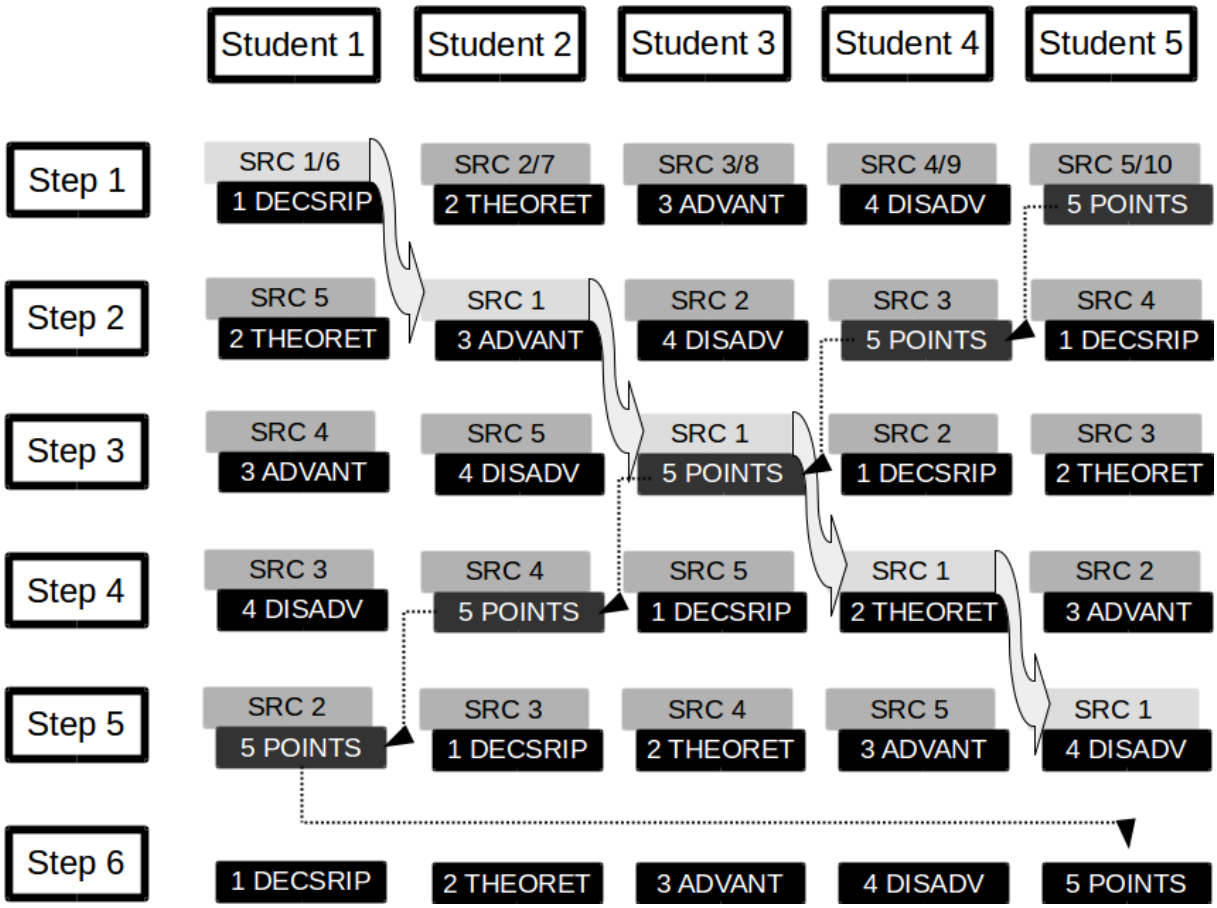

Figure 1. Overview of the six steps suggested in the collaboration script. The gray squares represent the sources (SRC) source. Source 1-5 were the main sources, source 6-10 were the additional sources. The black squares represent the wiki pages. Arrows show how students were supposed to shift sources and continue to work on different pages.

Table 1

Overview of the questions and the reliability of the coded categories within the answer of each question.

		initial alpha	definitive alpha
1. Definition	Please provide a description of peer assessment (i.e. what is peer assessment)		
1.1. Peers	Are peers or people with equal status mentioned?	0.79	0.94
1.2. Assessment	Is assessment or feedback mentioned?	0.79	0.86
1.3. Criteria	Is the use of criteria mentioned?	0.87	0.91
1.4. Learning	Is the value for learning mentioned?	0.75	0.82
1.5. Instructor	Is the role of the instructor mentioned?	0.58	0.83
2. Relation	How can peer assessment be related to other forms of assessment?		
2.1. Involvement	Is students involvement (responsibility) mentioned?	0.68	0.89
2.2. Instructor-assessment	Is the relation with instructor-assessment mentioned?	0.79	0.9
2.3. Co-assessment	Is the relation with co-assessment mentioned?	0.97	0.99
2.4. Self-assessment	Is the relation with self-assessment mentioned?	0.92	0.95
3. Background	What are the links between peer assessment and learning theories, learning principles, and/or assessment culture?		
3.1. Constructivism	Is constructivism mentioned?	0.98	0.99
3.2. Active role	Is the active role of students (responsibility, involvement, etc.) mentioned?	0.78	0.84
3.3. For learning	Is assessment for learning (tool for learning, shift from testing towards assessment, etc.) mentioned?	0.45	0.79
4. Conditions	What are the conditions for implementing peer assessment in an efficient way?		

4.1. Training	Is training (preparation of the students, etc.) mentioned?	0.63	0.87
4.2. Criteria	Is the use of criteria (standards, etc.) mentioned?	0.89	0.95
4.3. Objectivity	Is objectivity (fairness, reliability, etc.) mentioned?	0.86	0.91
4.4. Coaching	Is coaching (guidance, steering, etc.) by a teacher mentioned?	0.72	0.87
5. Added value	What is the added value of peer assessment for teachers and students?		
5.1. Involvement	Is higher involvement (ownership, responsibility, etc.) mentioned?	0.75	0.97
5.2. Efficiency	Is efficiency (reduction of time, effort, etc.) mentioned?	0.79	0.96
5.3. Fairness	Is fairness (objectivity) mentioned?	0.52	0.84
5.4. Development	Is personal development mentioned?	0.71	0.95
5.5. Metacognition	Are metacognitive skills mentioned?	0.47	0.86
5.6. Social skills	Are social skills (collaboration, interaction, communication, etc.) mentioned?	0.72	0.93
5.7. Academic skills	Are academic skills (critical thinking, analyzing, etc.) mentioned?	0.71	0.95
5.8. Performance	Is better performance (quality of learning) mentioned?	0.57	0.78

Note: all categories were coded binary (0 = not mentioned, 1 = mentioned) except for the categories 1.2., 1.3., and 4.1.

1.2. Assessment was coded 0 = not mentioned, 1 = assessment mentioned, 2 = feedback mentioned, 3 = assessment as well as feedback mentioned. This variable was scored 0, 0.5, 0.5, and 1 respectively.

1.3. Criteria was coded 0 = not mentioned, 1 = criteria mentioned without more, 2 = a priori available criteria mentioned, 3 = self-made criteria mentioned, 4 = both a priori available and self-made criteria mentioned. This variable was scored 0, 0.5, 0.75, 0.75, and 1 respectively.

4.1. Training was coded 0 = not mentioned, 1 = training mentioned but not specified, 2 = training for supporting skills (metacognitive, social) mentioned, 3 = training for peer-assessment (judging, giving feedback) mentioned, 4 = both forms of training mentioned. This variable was scored 0, 0.5, 0.75, 0.75, and 1 respectively.

For calculating the overall score on each of the five questions, the mean of the underlying categories was calculated.

Table 2

Model estimates for the two-level analysis of students' experienced quality of the collaboration, time invested, and number of sources read.

	Experienced (0-10 scale)		Time invested (in hours)		Sources read (count)	
	(a0)	(a1)	(b0)	(b1)	(c0)	(c1)
Fixed						
Intercept	7.738***	8.038***	13.609***	11.559***	7.819***	6.342***
	(0.202)	(0.273)	(0.947)	(1.226)	(0.367)	(0.388)
Scripted condition		-0.621^{ns}		4.190*		2.972***
		(0.393)		(1.758)		(0.551)
Random						
Group level variance	0.088	0.010	18.083*	13.366	4.372***	2.146**
	(0.392)	(0.374)	(8.011)	(6.962)	(1.176)	(0.667)
Student level var.	6.810***	6.793***	69.680***	69.882***	3.343***	3.35***
	(0.815)	(0.812)	(8.599)	(8.620)	(0.402)	(0.403)

Values between brackets are standard errors

*p < .05 **p < .01 ***p < .001 ^{ns}p = .115

Table 3

Average scores for Likert-scale items from 1 (completely disagree) to 7 (fully agree) for students in the scripted and the non-scripted condition.

	$M_{non\text{-}scripted}$[1] (SE)	Signif.[2]	$M_{scripted}$[3] (SE)
In our group, we divided the work [ns 4]	**6.26** (0.10)	>	5.89 (0.15)
I mostly wrote only one part of the wiki [**]	**5.03** (0.16)	>>>	2.76 (0.24)
Everybody was responsible for only one part of the wiki [ns]	**5.47** (0.19)	>>>	3.73 (0.27)
I only read one or two parts of the wiki [ns]	**2.02** (0.13)	>>>	1.33 (0.19)
I only felt responsible for one single part of the wiki [ns]	**2.55** (0.16)	>>	1.87 (0.23)
In our group, we tackled the work together [ns]	5.10 (0.15)	<<	**5.63** (0.21)
I helped writing most parts of the wiki [***]	3.02 (0.14)	<<<	**6.01** (0.20)
Everybody was responsible for all parts of the assignment [**]	3.45 (0.17)	<<<	**5.21** (0.25)
I read all parts of the wiki [ns]	5.63 (0.14)	<<<	**6.65** (0.20)
I felt responsible for all parts of the wiki [ns]	4.65 (0.17)	<<<	**5.69** (0.25)

[1] $M_{non\text{-}scripted}$ = Mean non-scripted condition (intercept)

[2] Signif. = Significance indication:

<, <<, and <<< significantly smaller than at respectively $p < .05$, $p < .01$, and $p < .001$

>, >>, and >>> significantly larger than at respectively $p < .05$, $p < .01$, and $p < .001$

[3] $M_{Scripted}$ = Mean scripted condition (i.e. intercept + scripted condition parameter)

[4] Significance of group variance is indicated in superscript at the end of each item, $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$. [ns] is indicating nonsignificance

Table 4

Multilevel analyses results of the log-file data

	(a) Number of edits	(b) Number of main pages edited (out of 5)	(c) Number of turn-takes on the five main pages
Fixed			
Intercept	50.938 *** (3.476)	3.015 *** (0.192)	4.032 *** (0.603)
Scripted condition	0.123 (5.012)	1.693 *** (0.275)	3.349 *** (0.859)
Time invested[1]	0.907 ** (0.319)	0.029 (0.015)	-0.004 (0.031)
Scripted*Time invested	-0.121 (0.470)	-0.023 (0.023)	0.098 * (0.048)
Random			
Group level variance	72.972 (55.285)	0.359 * (0.164)	5.607 *** (1.590)
Student level variance	685.711 *** (84.536)	1.478 *** (0.182)	5.605 *** (0.693)

Values between brackets are standard errors -- *p < .05 **p < .01 ***p < .001

[1]Time in hours, variable entered in model centered around mean

Table 5

Mean estimates for pre-test, post-test, and increase for the total score and for each of the five questions separately, based on three level models (group, student, measurement occasion).

	Total	Definition	Relation	Background	Conditions	Added value
		Overall (both conditions)				
M pre-test	12.3	25.6	7.4	7.8	12.5	8.0
M post-test	34.4	51.0	27.4	32.4	25.7	34.4
M increase	22.1 [a]	25.4 [a]	20.0 [a]	24.6 [a]	13.2 [a]	34.3 [a]
		Non-scripted condition				
M pre-test	13.0	26.0	8.5	7.7	12.9	9.7 [c]
M post-test	32.6	47.1	28.2	30.1	24.1	32.0
M increase	19.6 [a]	21.1 [a,b]	19.7 [a]	22.4 [a]	11.2 [a]	22.3 [a,d]
		Scripted condition				
M pre-test	11.3	25.2	6.2	8.0	12.0	6.4 [c]
M post-test	35.8	55.1	26.5	34.8	27.3	36.6
M increase	24.4 [a]	29.9 [a,b]	20.2 [a]	26.8 [a]	15.3 [a]	30.2 [a,d]

[a] There is a significant occasion effect ($p < .001$): post-test results are significantly higher

[b] There is a significant condition*occasion interaction effect ($p=.019$): scripted groups have a larger increase

c There is a significant condition effect ($p = .040$): at pre-test, scripted students score significantly lower

[d] There is a significant condition*occasion interaction effect ($p=.016$): scripted groups have a larger increase

Table 6

Two level models of post-test results for the total score and each of the five questions separately (controlled for time invested and sources read).

	Total	Definition	Relation	Background	Conditions	Added value
Fixed						
Intercept	33.772 ***	46.765 ***	32.129 ***	32.308 ***	23.592 ***	34.735 ***
	(1.814)	(2.132)	(3.432)	(3.767)	(3.320)	(2.452)
Pre-test score[1]	0.194	-0.121	0.198	-0.076	0.124	0.177
	(0.108)	(0.082)	(0.157)	(0.129)	(0.108)	(0.141)
Scripted condition	2.915	7.040 *	-6.050	-0.671	9.999	3.027
	(2.886)	(3.530)	(5.586)	(6.020)	(5.234)	(4.050)
Time invested[1]	0.080	0.145	0.201	-0.026	-0.022	0.246
	(0.165)	(0.202)	(0.327)	(0.326)	(0.279)	(0.230)
Scripted * Time inv.	0.324	-0.087	0.308	1.132 *	-0.125	0.314
	(0.236)	(0.294)	(0.467)	(0.478)	(0.414)	(0.333)
Sources read[1]	0.711	-0.476	1.850	1.883	-0.595	1.635 *
	(0.522)	(0.642)	(1.014)	(1.067)	(0.912)	(0.734)
Scripted * Sources	-1.770	0.784	-2.002	-1.618	-2.373	-2.963
	(1.149)	(1.463)	(2.289)	(2.368)	(2.020)	(1.648)
Random						
Group level variance	5.322	0.504	2.175	54.683	50.692	0.000
	(11.635)	(16.901)	(42.604)	(52.630)	(39.636)	(0.000)
Student level var.	167.377 ***	287.927 ***	703.798 ***	669.593 ***	472.976 ***	368.677 ***
	(21.732)	(35.791)	(88.841)	(85.410)	(60.065)	(40.346)

Values between brackets are standard errors -- *p < .05 **p < .01 ***p < .001

1 Variables entered in model centered around mean