Article

Structuring asynchronous discussion groups: The impact of role assignment and self-assessment on students' levels of knowledge construction through social negotiation

Authors

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M.

Status

Published in Journal of Computer Assisted Learning, Volume 25 Issue 2, Pages 177 - 188

Available:

DOI:

10.1111/j.1365-2729.2008.00292.x

http://dx.doi.org/10.1111/j.1365-2729.2008.00292.x

Please cite as:

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2009). Structuring asynchronous discussion groups: The impact of role support and self-assessment on students' levels of knowledge construction through social negotiation. *Journal of Computer Assisted Learning, 25*, 177-188.

Homepages authors:

http://www.onderwijskunde.ugent.be/en/cv_dewever.htm

http://www.onderwijskunde.ugent.be/en/cv_vankeer.htm

http://www.onderwijskunde.ugent.be/en/cv_schellens.htm

http://www.onderwijskunde.ugent.be/en/cv_valcke.htm

Related CSCL articles:

Structuring asynchronous discussion groups: The impact of role assignment and self-assessment on students' levels of knowledge construction through social negotiation

Bram De Wever[ab], Hilde Van Keer[a], Tammy Schellens[a], and Martin Valcke[a]

[a]Department of Education, Faculty of Psychology and Educational Sciences, Ghent University, H. Dunantlaan 2, B-9000 Gent, Belgium
[b]Postdoctoral research fellow of the Research Foundation Flanders (FWO)

Abstract

This article examines the impact of the introduction of roles and the added value of self-assessment on students' level of knowledge construction in online asynchronous discussions in a first year university course in instructional sciences.

Students' postings in 20 discussion groups were used as the research data for this study. All messages, submitted during the 12 week discussion period, comprising 4 discussion themes of 3 weeks each, were analysed. Repeated measures multilevel modelling was adopted to analyse the data of the content analysis.

The results point at a significant positive impact of assigning roles to students. However, this positive impact depends on the moment of the introduction the roles. Higher levels of social knowledge construction were found in discussion groups where roles were introduced right at the start of the discussions and faded out towards the end. The results further indicate that self-assessment has no significant added value.

Keywords: computer supported collaborative learning (CSCL), collaboration, role assignment, self-assessment, scripting, networked learning

Introduction

In the early days of the information technology age, computers were rather positioned as personal tools and their potential to foster interpersonal communication was less well anticipated (Crook, 2002). In contrast, current approaches towards computers and the Internet acknowledge this interpersonal significance. Recent online learning and instruction approaches highlight the importance of learner interaction in view of knowledge construction. This has resulted in a growing implementation of computer-supported collaborative learning (CSCL) approaches, including asynchronous discussion groups, whose educational potential is often studied (e.g. Järvelä & Häkkinen 2002; De Laat & Lally 2004; Schellens & Valcke 2005; Schellens *et al.* 2005).

This study focuses on two instructional approaches to stimulate knowledge construction through social negotiation in asynchronous e-discussions, namely role assignment and self-assessment. The study is situated in the context of a first year instructional sciences course, where asynchronous discussion groups of 10 students are organised in addition to weekly face-to-face working sessions. The discussion groups are organised to foster students' processing of the learning content and, by confronting them with authentic tasks, to promote discussion about the different concepts presented in the face-to-face sessions and the course manual. Roles were assigned to students when collaborating in the online asynchronous discussions in order to promote knowledge construction through social negotiation. Previous research presented empirical evidence that students act in line with assigned roles (De Wever *et al.* in press). This specific structuring approach is combined with the introduction of self-assessment in order to enhance reflection. The main aim of this study is to evaluate the impact of role assignment on students' knowledge construction and to study the surplus value of introducing self-assessment.

Roles as structuring tool

Putting individual students together does not necessarily bring about effective interaction or collaborative learning (Weinberger *et al.* 2005). Instructional design, building on collaborative learning environments, therefore focuses on embedding a certain amount of structure, such as adding specific goals, defining task types, presenting task prescripts, or pre-structuring (scripting) (De Wever *et al.* in press). The goal of introducing structure is to support interaction processes and actual collaborative learning within CSCL-environments. Some empirical evidence underscores that pre-structuring or scripting learning environments improves collaboration (Pfister & Mühlpfordt 2002) and produces strong positive learning effects (Baker & Lund 1997).

In a previous article (De Wever *et al.* in press) we examined role assignment as a specific type of scripting. Five roles were presented to students: the role of starter, moderator, theoretician, source searcher, and summariser. The study showed that students act in line with their assigned roles. Due to this manipulation check, the present study can in a valid way assume that possible treatment effects can be related to differential role assignment and adoption.

Self-assessment to enhance reflection

Since McLoughlin and Luca (2002) argue that CSCL-environments enable students to become more self-directed, and that "the shift to student self-direction and autonomy means that students need to take more responsibility for their own learning" (p. 577), self-assessment was introduced in this study. Self-assessment requires learners to make judgements about their own learning and is considered as a tool providing feedback to students about both learning and educational standards (Boud 1995; Boud & Falchikov 1989). It requires students to consider the characteristics of competent work in a given area or situation, and to apply these criteria to their own work (Boud 1999). Self-assessment helps students to internalise academic standards (Gibbs 2006). As such, self-assessment encourages independent and self-directed learning. In collaborative contexts, this implies that self-assessment fosters reflection on the quality of personal contributions and the input of others, and to develop awareness of effective and qualitative contributions to the discussions (Sluijsmans *et al.* 1999; Freeman & McKenzie 2002). Students need to monitor the actual condition of their discussion, learning process, and human relations, in order to improve their learning community and to plan their upcoming study so that they should make their learning substantial (Mochizuki *et al.* 2003).

While performing their own regular and structured self-assessment, learners develop a questioning and reflective approach (Robinson & Udall 2006). Self-assessment encourages students to become critical and perceptive, stimulates reflection, and is thereby contributing to the learning processes and to lifelong learning (Larres *et al.* 2003). Empirical evidence stresses that self-assessment has an effect on cognition, affection, and conation and can encourage deep approaches to learning (McDonald & Boud 2003). Research also reveals a considerable impact of self-assessment on students' content-related learning, quality of problem solving, and self-reflection (Sluijsmans *et al.* 1999).

In this respect, self-assessment was introduced in the present study as a reflection tool and a tool for learning. It was implemented primarily as a way to help students improve their learning, as it focuses students' attention on the metacognitive aspects of their learning and teaches them to be more effective at monitoring their own performance (Longhurst & Norton 1997), and not as a substitute for the instructor's evaluation. Following the claim that self-assessment is clearly an important part of supporting students to improve their own learning (Longhurst & Norton 1997), it is hypothesised that self-assessment of the individual

contributions in a CSCL-environment will elicit readjustment of discourse in forthcoming collaborative activities. The idea is that by asking students to reflect upon and to rate the quality of their performance, students will identify weaknesses and strengths and might amend or redirect their contributions in forthcoming discussions (Hunt *et al.* 2002).

Method

Research questions

The first research question focuses on whether or not students are capable of judging their own social knowledge construction processes. Since self-assessment has a considerable impact on self-reflection (Sluijsmans *et al.* 1999) and reflecting on one's own knowledge construction processes might influence the quality of the knowledge construction processes of subsequent discussions, we want to check to what extent students are able to assess their own knowledge construction processes accurately. This question precedes the study of the impact of the research conditions on the knowledge construction processes.

The second research question focuses on (1) determining whether role assignment has an impact on the knowledge construction processes in the discussion groups, (2) whether the moment of introduction of the role assignment is an important factor, and (3) whether self-assessment has a surplus value to stimulate students' knowledge construction through social negotiation.

Research setting

Context

The present study was conducted in the context of a first year course in instructional sciences in the bachelor in Educational Sciences of Ghent University. The instructional design of this course combined face-to-face sessions with an online learning environment (Schellens & Valcke 2000). All first year students enrolled (N = 273) participated in the discussion groups.

The discussion groups were organised in parallel to the weekly face-to-face sessions to promote the timely study of the theoretical concepts. It was expected that students would develop a stronger knowledge base when applying the theoretical concepts during discussions and while they were involved in social negotiations and debate. After a one-week trial discussion, the formal study plan required students to discuss four successive authentic tasks. Each discussion took three weeks. Within the three-week periods, students collaborated online, independently of time and location. The authentic discussion tasks were identical for all groups and were related to corresponding chapters in the handbook (behaviourism, cognitivism, constructivism, and evaluation). The main goal was to stimulate debate on the theoretical concepts presented in the face-to-face sessions and course manual. A full example of one of the discussion tasks can be found in the supplementary material.

Students were divided at random into discussion groups of 10 persons. Participation in the online discussion groups was a formal component of the course and represented 25% of the course grade. Students were required to contribute at least four times per discussion theme. In line with the approach of Guldberg and Pilkington (2006), "the tutor monitors the discussion but takes a background role unless a real need to intervene is perceived" (p. 160).

Roles

In specific research conditions, particular students of a group were assigned one of the following five roles: starter, summariser, moderator, theoretician, and source searcher. The inclusion of the starter and summariser was founded on the literature regarding the starter-wrapper technique (Hara, Bonk, & Angeli, 2000), while the moderator was incorporated based on the findings of Gray (2004) indicating the role of an online moderator as critical for enhancing learning. Further, Strijbos et al. (2004) argue that when cooperative learning pedagogies, and more specifically roles, are used in higher education or online learning environments, they should be adapted to the specific context, as students in these settings vary considerably in prior knowledge, experience, and collaboration skills. Taking into account that the discussion groups are organised in order to stimulate debate on the theoretical concepts presented in the face-to-face sessions and course manual, the starter, summariser, and moderator role were supplemented with the role of source searcher and theoretician.

The *starter* was required to start off the discussion, to add new points for other students to build upon, and to give new impulses when discussions slacked off. The role of the *moderator* consisted of monitoring the discussion, asking critical questions, and probing others' opinion. Students in the role of *theoretician* were required to introduce theoretical information and to ensure that all relevant theoretical concepts were used in the discussion. The role of the *source searcher* comprised of seeking external information about the discussion topics in order to stimulate other students to go beyond the scope of the available handbook. The *summariser* was expected to post interim summaries during the discussion and a final synopsis at the end, focusing on identifying dissonance and harmony between the messages and drawing conclusions. In general, all students were allowed to perform all these activities. However, students with a specific role were asked to pay explicit attention to the activities related to their role on a very regular basis.

Self-assessment

In the present study, self-assessment was introduced to enhance reflection and to stimulate self-directed learning (Larres *et al.* 2003). The students were asked to evaluate themselves in relation to the knowledge construction processes in their messages. They were informed by the staff members about the fact that this self-assessment would not affect the formal score for this course and about the criteria for the summative assessment. The self-assessment was based on an online questionnaire in which students had to rate their knowledge construction through social negotiation after each discussion assignment.

Data collection

The discussion transcripts of 20 discussion groups were selected for this study and the transcripts of the entire 12 week discussion period were analysed, comprising 4 discussion themes of 3 weeks each. This resulted in the analysis of 4818 messages or approximately 60453 lines.

Design

Discussion groups were assigned to one of three research conditions. In condition 1, students started discussing without role assignment in theme 1 and 2; role assignment was introduced when discussing theme 3 and 4. In condition 2, roles were assigned right from the start in theme 1 and 2 but no longer stressed during theme 3 and 4. The third condition was

equal to the second condition, except that students in the third condition were in addition requested to fill out a self-assessment questionnaire at the end of each discussion theme.

The specific cross-over design of the present study was helpful to answer research question 2, since it allows us to explore the differences between role-supported and non-role-supported discussions. Furthermore, the comparison of the first two research conditions enables us to study whether or not the timing of role assignment is an important mediating factor influencing students' knowledge construction through social negotiation. Comparing the second and the third condition allows us to explore whether or not self-assessment has a surplus value in stimulating knowledge construction through social negotiation. Table 1 gives an overview of the different research conditions.

Table 1. Overview of research conditions

	Condition 1	Condition 2	Condition 3
Theme 1	No Role Assignment	Role Assignment	Role Assignment + SA
Theme 2	No Role Assignment	Role Assignment	Role Assignment + SA
Theme 3	Role Assignment	No Role Assignment	No Role Assignment + SA
Theme 4	Role Assignment	No Role Assignment	No Role Assignment + SA

SA = Self-assessment

In each condition, roles were introduced in either the first or the last two discussion themes. In the first theme where role assignment was applied, five randomly selected students were given one of the five roles. In the second discussion theme with role assignment, the roles were assigned to the students who did not take up a role in the first discussion theme. The rotation of roles guaranteed that each student adopted a specific role at least once. Students were asked to perform their roles in addition to submitting regular discussion input. Taking into account the different discussion assignments, the study was constructed according to a repeated-measures design.

Quantitative content analysis

In order to determine the level of knowledge construction through social negotiation, quantitative content analysis was applied. This quantitative approach to content analysis focuses on collecting and coding a large amount of data. This large dataset of codes allows us to perform statistical tests to compare the different conditions.

The interaction analysis model of Gunawardena *et al.* (1997) was applied to analyse the transcripts. This model examines the social construction of knowledge in computer conferencing and distinguishes five different levels of knowledge construction activities: (1) sharing and comparing information, (2) identifying areas of disagreement, (3) negotiating meaning and co-construction of knowledge, (4) evaluation and modification of new schemas that result from co-construction, and (5) reaching and stating agreement and application of co-constructed knowledge (See Table 2). A detailed discussion of this model can be found in De Wever *et al.* (2006). This model of Gunawardena *et al.* (1997) has been applied in a number of empirical studies (Marra *et al.* 2004; Schellens & Valcke 2005; Schellens *et al.* 2005; De Wever *et al.* 2006). Marra *et al.* (2004) compared this model with the model of Newman *et al.* (1995) and argued that the Gunawardena's model provides "a more holistic view of discussion flow and knowledge construction" (p. 39). Research of Schellens and Valcke (2005) studied the validity of the instrument of Gunawardena *et al*. (1997) by simultaneously coding the discussions using the instrument of Veerman and Veldhuis-Diermanse (2001).

They concluded that both models are parallel to one another for the first three levels of knowledge construction. They furthermore conclude that the coding scheme of Gunawardena *et al*. (1997) goes beyond the scheme of Veerman and Veldhuis-Diermanse (2001) and discriminates more advanced levels of knowledge construction, such as testing and applying newly constructed mental models.

As suggested by Rourke *et al.* (2001), messages were selected as units of analysis since complete messages are an objective unit and are considered as the unit defined by the original author of the contributions.

A discussion of the analysis scheme of Gunawardena *et al*. (1997) can be found in the supplementary material, together with a discussion of coding decisions and an excerpt of student postings with the assigned codes.

Table 2.
Levels of knowledge construction in the interaction analysis scheme of Gunawardena et al. (1997)

Level	Description
1	Sharing and comparing of information
2	Exploration of dissonance
3	Negotiation of meaning
4	Testing synthesis
5	Agreement statements and applications

By analogy with the content analysis scheme applied to analyse the transcripts of the discussion groups, the self-assessment questionnaire was founded on the instrument of Gunawardena *et al.* (1997). The questions probe into students' perceptions of their achieved levels of knowledge construction through social negotiation. More particularly, students were asked to rate how often their own contributions to the discussion fit into each of the five levels of knowledge construction. An example of the self-assessment items was "My contributions aimed at sharing and comparing of information".

By presenting the self-assessment questionnaire after each discussion, students were required to step back and evaluate the levels of knowledge construction reflected in their contributions. They were encouraged to reflect on the extent to which their messages were effective contributions to the ongoing discussion. In this way, students were required to monitor their discussion behaviour. They were motivated to verify which knowledge construction processes they invoked. In case they noticed their messages did not cover the whole spectrum of knowledge construction processes, this could lead them into adjusting their future discussion behaviour in order to optimise future debates.

Coding strategy and reliability

Five independent coders were trained during approximately 7 hours to carry out the coding activity. After working with coding examples for each level of knowledge construction in the analysis model (Gunawardena *et al.* 1997), they coded some transcripts together in order to discuss and elaborate on the coding process. Next, the transcripts were coded independently. A number of transcripts were randomly selected for calculating interrater reliability coefficients. The Krippendorff's alpha interrater reliability coefficient ($\alpha = .52$, $n = 198$) was situated between .40 and .80, which corresponds to 'fair to good agreement beyond chance' (Banerjee *et al*. 1999; Neuendorf 2002; De Wever *et al.* 2006).

Statistical analysis

Taking into account the hierarchical nesting of students in discussion groups and the successive nature of the four themes, repeated-measures multilevel modelling was applied in order to answer the research questions. In the present study, measurement occasions (the four discussion themes) are nested within subjects (Hox 1998). We refer to De Wever *et al.* (2007) for an in-depth discussion on this analysis technique.

The statistical package R 1.8.1. was used for the calculation of the interrater reliability. MLwiN 2.01. was used to perform the multilevel analysis. The multilevel models were estimated with the iterative generalised least squares (IGLS) procedure. All analyses assume a 95% confidence interval.

Results

Research question 1: Students' ability to evaluate their own social knowledge construction processes

In order to explore whether students are able to assess their own level of knowledge construction, we focus on the match between students' self-assessment and the content analysis of their messages. Following Longhurst and Norton (1997) a convergence measure was computed per discussion theme by calculating the difference between the self-reported occurrence of utterances reflecting each level of knowledge construction (LKC_{SA}) and the observed occurrence of messages for each level of knowledge construction (LKC_{OBS}) as coded by the coders during the content analysis.

For each level of knowledge construction, a difference score was calculated (LKC_{DIF} 1 to 5). Negative difference scores indicate that students underestimate their level of knowledge construction, while positive scores point at overestimation. The multilevel analyses indicate that students underestimate themselves at the first level of knowledge construction (LKC_{DIF} 1 = -1.103, *SE* = 0.128) and overestimate themselves at the four subsequent levels (LKC_{DIF} 2 = 1.314, *SE* = 0.086; LKC_{DIF} 3 = 1.344, *SE* = 0.101; LKC_{DIF} 4 = 2.280, *SE* = 0.084; LKC_{DIF} 5 = 1.714, *SE* = 0.067). A detailed discussion of these results can be found in the supplementary material.

Research question 2: Impact of the research conditions on levels of knowledge construction reflected in the online discussions

The second research question focuses on the impact of the three different research conditions on the levels of knowledge construction reflected in student messages (LKC_{OBS}). For this research question, a four-level model was estimated, with messages (level 1) hierarchically nested within measurement occasions (level 2) that are clustered within students (level 3) who are in turn assigned to groups (level 4). The analysis models were built following a stepwise procedure. A random intercept null model and a compound symmetry model were estimated first. Next, additional analyses were performed in which the different research conditions were included as predictors to the model. All models are presented in the supplementary material.

In the final model (see Table 3) the difference in LKC between the four themes and the two research conditions is explored in detail taking the interaction effects between the conditions and the themes into account. In this respect the differential progress in LKC_{OBS} in the different research conditions is studied. In the first theme students' contributions reflect

significantly higher LKC_{OBS} in both condition 2 (role/no-role) and condition 3 (role/no-role+SA) compared to the first condition (no-role/role) (respectively $\chi^2 = 11.725$, $df = 1$, $p = .001$ and $\chi^2 = 5.767$, $df = 1$, $p = .016$). No significant differences were found between condition 2 and 3 in theme 1 ($\chi^2 = 1.128$, $df = 1$, $p = .228$).

Table 3. Multilevel parameter estimates of LKC_{OBS} (Final Model)

Parameter			
Fixed		*Random*	
Intercept	1.244 (0.065)	Level 4 – group	
Theme 2 (cognitivism)	-0.058 (0.066)	$\sigma^2 f0$	0.013 (0.005)*
Theme 3 (constructivism)	0.070 (0.061)	Level 3 – student	
Theme 4 (evaluation)	0.040 (0.061)	$\sigma^2 v0$	0.004 (0.004)
Condition 2	0.297 (0.087)***	Level 2 – theme	
Condition 3	0.210 (0.088)*	$\sigma^2 u0$	0.006 (0.006)
Theme 2 * Condition 2	-0.150 (0.088)	Level 1 – message	
Theme 2 * Condition 3	-0.064 (0.088)	$\sigma^2 e0$	0.617 (0.013)***
Theme 3 * Condition 2	0.444 (0.081)***		
Theme 3 * Condition 3	0.228 (0.081)**		
Theme 4 * Condition 2	-0.056 (0.084)		
Theme 4 * Condition 3	0.045 (0.084)		

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

In the third theme, the LKC_{OBS} is also significantly higher in condition 2 (role/no-role) and condition 3 (role/no-role+SA) compared to the first condition (no-role/role) (respectively $\chi^2 = 29.824$, $df = 1$, $p < .001$ and $\chi^2 = 7.853$, $df = 1$, $p = .005$). In addition, the LKC_{OBS} is significantly higher in condition 2 compared to condition 3 ($\chi^2 = 7.954$, $df = 1$, $p = .005$). The differences between the conditions are not significant for themes 2 and 4. As to the differences between role and no role based discussions, the results of model 3 indicate that for the initial discussion theme the discussion groups with role assignment reach higher levels of knowledge construction, whereas the opposite is true for theme 3. In condition 2 and 3 no role assignment was present in this third discussion theme and yet these discussion groups reach higher levels of knowledge construction.

Discussion

Research question 1

The first question focused on the ability of freshmen to evaluate their level of knowledge construction in an accurate and critical way.

The results clearly indicate that students underestimate the extent to which they engage in sharing and comparing information during the ongoing discussion. On the other hand, they overestimate the occurrence of postings reflecting the four subsequent levels of knowledge construction. This means that students post fewer contributions focusing on identifying disagreement, negotiating meaning, evaluating co-constructed meaning, and agreeing on and applying the co-constructed knowledge than they actually think they do. These results are in line with Robinson and Udall's statement (2006) that "students are often unable to make realistic judgements about their own learning" (p. 98).

The findings that first-year students are not always capable of judging themselves accurately might be due to lack of experience. This corresponds to the findings of Larres *et al.* (2003), who studied the difference between objective and self-appraisal computer literacy tests and argued that at entry level students "would require much more experience in self-evaluation before it to become effective" (p. 109). The findings can entail that more support

should be given to the students to develop their self-assessment skills; for example by making students aware of the fact that their self-assessed ratings will be validated with ratings from other sources (e.g. cross-checking with other measures or verification with peer or instructor assessment) and by providing comparative information about peers as suggested by Larres *et al.* (2003) and Gibbs and Simpson (2004). In addition, students could be informed of the divergence in self-assessed ratings and independent ratings. Finally, explicit development of assessment skills can be called for, as suggested by McLoughlin and Luca (2002) and Black *et al.* (2004) who argue that students might need assistance in achieving the skills that come with more autonomy and responsibility.

Falchikov and Boud (1989) point at two other possible explanations for the lack of students' accuracy in self-assessment. First, they claim that "studies within the broad area of science appear to produce more accurate self-assessment than do those from other areas of study" (p. 425). In addition, they claim that the level of the course of which the assessment is a part of, is an important influential factor as well. It more particularly appears that students in advanced courses are better at assessing themselves accurately.

Based on the findings with respect to the first research question, it can be argued that future practice and research should aim at making students' self-assessment more accurate by exposing them to self-assessment more frequently, by offering an introductory training, by making students aware of the fact that their self-assessment will be monitored, and by providing them with comparative information and feedback.

Research question 2

The second research question focused on the impact of (1) role assignment, (2) the moment of introducing the roles, and (3) the surplus value of self-assessment.

When we focus on the results with regard to the introduction of role assignment, significant differences were found in theme 1 and 3 between the *condition with roles in the two final discussion themes* and the *conditions with roles during the two initial themes*. In both theme 1 and 3, the latter conditions outperform the former one with respect to the levels of knowledge construction. As to the impact of the presence of roles, this implies that in the first theme students in discussion groups with roles outperform students discussing in groups without role assignment. In the third theme the opposite can be concluded. Concerning the importance of the moment of the role introduction, it can be noticed that in both the first and the third theme, groups with initial role assignment outperform groups receiving role assignment at the end even when the original role assignment had faded out.

These results lead us to the conclusion that the moment of time of the role introduction can have an important impact. The observation that groups with initial role assignment outperform the others in theme 3 might point at the fact that students have interiorised the role-related activities. In this respect, Weinberger *et al.* (2005) argue that "fading of the cooperation script could improve internalization processes". However, since the trend is not pursued in the fourth theme, further research is needed to confirm this finding. Further research might also focus on gradually decreasing the role assignment, since Hoadley and Enyedy (1999) argue that "we know from studies of learning technology that gradually fading of scaffolding from a tool, or tools with a gradually sloped learning curve are more effective than sudden drops in scaffolding, or tools with a staircase shaped learning curve" (p. 250).

With regard to the impact of self-assessment on students' knowledge construction, the research condition without self-assessment significantly outperforms its equivalent including self-assessment in the third discussion theme. From these findings, the conclusion can be drawn that the introduction of recurrent self-assessment procedures does not have a

significant surplus value on knowledge construction processes in the asynchronous discussions.

This may be due to the fact that the first-year students in our study were not yet able to assess their knowledge construction processes in an accurate way. In this respect, the introduction of training students in self-assessment merits particular attention in future research since McDonald and Boud (2003) already illustrated that "self-assessment training had a significant impact on the performance of those who had been exposed to it" (p. 217).

Moreover, further research is wanting since it can be argued that the process of incorporating self-assessment to enhance the quality of the discourse in online discussion groups is still in its infancy. Murphy and Jerome (2005) note that "little has been written on students' self-assessment of participation in online discussion". In this respect, they suggest the use of self-analysis comprising a detailed examination of the number of messages, their distribution over the modules, and their length, supplemented with an analysis of the content of the contributions in relation to claims and grounds and a critical assessment to demonstrate knowledge construction by presenting quotes. Such a detailed self-analysis might have a more direct impact on knowledge construction in discussion groups. However, further research is needed to confirm this hypothesis. In addition, future studies should also focus more on the long term effects, since students may need more time and experience in self-assessment in order to improve their participation in the discussion groups.

Limitations of this study and directions for future research

Since the present study aims to study collaborative learning in CSCL-environments by manipulating variables that influence collaborative activities (see O'Donnell & Dansereau 1992), we focus on the social knowledge construction processes in the discussions. Studying processes is important, "especially if educators want to know which learning activities and methods are contributing to collaborative knowledge building" (Dennen & Paulus 2006, p. 1). In online discussions, it is therefore necessary to look at what is actually going on during students' discourse (Schellens *et al.* 2007). Consequently, the present research studies a process-related dependent variable that is an indicator of knowledge construction in the online discussions (Dennen & Paulus 2006). Further research should focus on unravelling the specific relationship between knowledge construction processes and the actual acquisition of knowledge, for instance by presenting knowledge acquisition tests after each discussion theme.

Another limitation of this study is the sole use of quantitative content analysis. This technique allows us to focus on one aspect and explore this meticulously over a large amount of data. However, expanding this technique with a more qualitative approach to content analysis would help us to get a deeper insight in students' discourse and to pick up nuances and details of specific social knowledge construction processes within selected student postings. Moreover, our quantitative content analysis was based on a single coding scheme from Gunawardena *et al.* (1997). Although this model has been compared with other models and we are confident that this model analyses knowledge construction in a reliable way, there still remain two drawbacks. First, social knowledge construction can be operationalised in different ways. By selecting this model we were bound to this specific operationalisation. Other models may shed a different light on knowledge construction. A second drawback is that measuring knowledge construction is never completely accurate. Although we used multiple coders and paid specific attention to the coder training, a certain amount of indistinctness still remains. In order to overcome the above-mentioned drawbacks, future research could validate our findings by applying one or more alternative content analysis scheme(s).

Furthermore, this study took place in an authentic educational setting. This implies that we could not control all variables affecting instructional processes and outcomes. This is a limitation compared with experimental studies. However, it also presents advantages, since this complex and ecologically valid setting provides a more stringent test of the successful implementation of roles and self-assessment as compared to studies in controlled laboratory settings. In this respect, we argue that the interventions implemented in this study are feasible and that the results can be generalised to our research context, which is the study of knowledge construction processes and the related outcomes in online asynchronous discussion groups with first year university students. Further research, implemented in other knowledge domains and with students of different educational levels, is however needed in order to make more general statements about the impact of roles and self-assessment.

Conclusion

The main aim of this study was to examine the effect of assigning roles on students' knowledge construction and to study the surplus value of introducing reflection through self-assessment on the knowledge construction processes within the discussion groups.

With respect to the introduction of roles, it can be concluded that introducing roles is a valuable structuring tool, especially if roles are introduced at the start of the discussions and faded out at a later stage. In this respect, it appears that role assignment is particularly helpful to get students started. The ultimate goal is that this structuring tool eventually can be faded out or taken away when students have interiorised the skills related to the different roles and are competent enough to discuss in a more natural way, which is without the additional support and structure of role assignment. In this respect, we agree with Brown et al. (1989), who state that fading of support should be an integral part of scaffolding. The findings of the present study suggest that students were already sufficiently competent to move forward without the additional structure offered by explicit role assignment after discussing for six weeks.

As to the implementation of self-assessment, it can be concluded that a larger investment in support for the students should be made in order to increase freshmen's ability to assess their knowledge construction processes in asynchronous discussions accurately. This can be achieved by exposing them more frequently to self-assessment experiences, by implementing a self-assessment training, by pointing at the validation of their self-assessment and providing comparative information, by providing intermediate feedback by instructors or by peers, or by introducing peer assessment. As to the impact of self-assessment, this study failed to show a significant surplus value of self-assessment on the levels of knowledge construction reflected in students' discourse in asynchronous discussion groups. However, further research and practice is recommended since the students in this study were not experienced in assessing their knowledge construction processes, and research on incorporating self-assessment to enhance the quality of the discourse in online discussion groups is still in its infancy (Murphy & Jerome 2005).

References

Baker, M. & Lund, K. (1997) Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning **13***, 175-193.

Banerjee, M., Capozzoli, M., McSweeney, L. & Sinha, D. (1999) Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics **27***, 3-23.

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2004) Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan **86***, 9-21.

Boud, D. (1995) *Enhancing learning through self assessment*. Kogan Page, London.

Boud, D. (1999) Avoiding the traps: seeking good practice in the use of self assessment and reflection in professional courses. *Social Work Education* ***18***, 121-132.

Boud, D. & Falchikov, N. (1989) Quantitative studies of self-assessment in higher education: A critical analysis of findings. *Higher Education* ***18***, 529-549.

Brown, J. S., Collins, A. & Duguid, P. (1989) Situated Cognition and the Culture of Learning. *Educational Researcher* ***18***, 32-42.

Crook, C. (2002) Deferring to resources: collaborations around traditional vs computer-based notes. *Journal of Computer Assisted Learning* ***18***, 64-76.

De Laat, M. & Lally, V. (2004) It's not so easy: Researching the complexity of emergent participant roles and awareness in asynchronous networked learning discussions. *Journal of Computer Assisted Learning* ***20***, 165-171.

De Wever, B., Schellens, T., Valcke, M. & Van Keer, H. (2006) Content analysis schemes to analyse transcripts of online asynchronous discussion groups: A review. *Computers & Education* ***46***, 6-28.

De Wever, B., Schellens, T., Van Keer, H., & Valcke, M. (in press). Structuring asynchronous discussion groups by introducing roles: Do students act in line with assigned roles? *Small Group Research*. De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2007). Applying multilevel modelling on content analysis data: Methodological issues in the study of the impact of role assignment in asynchronous discussion groups. *Learning and Instruction* ***17***, 436-447.

Dennen, V. P. & Paulus, T. M. (2006) Researching "collaborative knowledge building" in formal distance learning environments. In *Computer supported collaborative learning 2005: The next 10 Years!* (eds. T. Koschman, T. W. Chan & D. D. Suthers). International Society of the Learning Sciences, Taipei. Retrieved September 1, 2006, from http://css.cscl2005.org/ Fullpapers.aspx

Falchikov, N. & Boud, D. (1989) Student self-assessment in higher education: A meta-analysis. *Review of Educational Research* ***59***, 395-430.

Freeman, M. & McKenzie, J. (2002) SPARK, A Confidential Web-Based Template for Self and Peer Assessment of Student Teamwork: Benefits of Evaluating across Different Subjects. *British Journal of Educational Technology* ***33***, 551-569.

Gibbs, G. (2006) How assessment frames student learning. In *Innovative assessment in higher education* (eds. C. Bryan & K. Clegg), pp. 23-36. Routledge, London.

Gibbs, G. & Simpson, C. (2004) Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education* ***1***, 3-31.

Gray, B. (2004). Informal learning in an online community of practice. *Journal of Distance Education* **19**, 20-35.

Guldberg, K. & Pilkington, R. (2006). A community of practice approach to the development of non-traditional learners through networked learning. *Journal of Computer Assisted Learning* ***22***, 159-171.

Gunawardena, C. N., Lowe, C. A. & Anderson, T. (1997) Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research* ***17***, 397-431.

Hara, N., Bonk, C. J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science* **28**, 115-152.

Hoadley, C. & Enyedy, N. (1999) Between information and communication: Middle spaces in computer media for learning. In *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference* (eds. C. Hoadley & J. Roschelle), pp. 242-251.

Lawrence Erlbaum, Mahwah NJ. Retrieved September 1, 2006, from http://www.tophe.net/papers/Hoadley-Enyedy-1999.pdf

Hox, J. J. (1998) Multilevel modeling: When and why. In *Classification, data analysis, and data highways* (eds. R. Balderjahn, R. Mathar & M. Schader), pp. 147-154. Springer-Verlag, New York.

Hunt, N., Hughes, J. & Rowe, G. (2002) Formative Automated Computer Testing (FACT). *British Journal of Educational Technology **33***, 525-535.

Järvelä, S. & Häkkinen, P. (2002) Web-based cases in teaching and learning: The quality of discussions and a stage of perspective taking in asynchronous communication. *Interactive Learning Environments **10***, 1-22.

Kollar, I., Fischer, F. & Hesse, F. W. (2003) Cooperation scripts for computer-supported collaborative learning. In *Proceedings of the International Conference on Computer Support for Collaborative Learning, CSCL 2003 - Community events, communication, and interaction* (eds. B. Wasson, R. Baggetun, U. Hoppe & S. Ludvigsen), pp. 59-61. Intermedia, Bergen.

Larres, P. M., Ballantine, J. A. & Whittington, M. (2003) Evaluating the validity of self-assessment: Mearsuring computer literacty among entry-level undergraduates within accounting degree programmes at two UK universities. *Accounting Education **12***, 97-112.

Longhurst, N. & Norton, L. S. (1997) Self-assessment in coursework essays. *Studies In Educational Evaluation **23***, 319-330.

Marra, R. M., Moore, J. L. & Klimczak A. K. (2004) Content analysis of online discussion forums: A comparative analysis of protocols. *Educational Technology Research and Development **52***, 23-40.

McDonald, B. & Boud, D. (2003) The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations. *Assessment in Education **10***, 209-220.

McLoughlin, C. & Luca, J. (2002) A learner-centred approach to developing team skills through web-based learning and assessment. *British Journal of Educational Technology **33***, 571-582.

Mochizuki, T., Fujitani, S., Isshiki, Y., Yamauchi, Y. & Kato, H. (2003) Assessment of Collaborative Learning for Students: Making the State of Discussion Visible for their Reflection by Text Mining of Electronic Forums. In *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003* (ed. G. Richards), pp. 285-288. AACE, Chesapeake, VA.

Murphy, E. & Jerome, T. (2005) Assessing students' contributions to online asynchronous discussions in university-level courses. *Electronic Journal of Instructional Science and Technology, **8***. Retrieved September 1, 2006, from http://www.usq.edu.au/electpub/e-jist/docs/vol8_no1/commentary/ stu_contrib_ansynch.htm

Neuendorf, K. A. (2002) *The Content Analysis Guidebook*. Sage Publications, Thousand Oaks, CA.

Newman, D. R., Webb, B., & Cochrane, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology **3***, 56-77. Retrieved September 1, 2007, from http://www.qub.ac.uk/mgt/papers/methods/contpap.html

O'Donnell, A. M. & Dansereau, D. F. (1992) Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In *Interaction in cooperative groups. The theoretical anatomy of group learning* (eds. R. Hertz-Lazarowitz & N. Miller), pp. 120-141. Cambridge University Press, New York.

Pfister, H. R. & Mühlpfordt, M. (2002) Supporting discourse in a synchronous learning environment: The learning protocol approach. In *Computer support for collaborative*

learning: Foundations for a CSCL community. Proceedings of CSCL 2002 (ed. G. Stahl). Lawrence Erlbaum, Mahwah, NJ. Retrieved September 1, 2006, from http://newmedia.colorado.edu/cscl/178.pdf

Robinson, A. & Udall, M. (2006) Using formative assessment to improve student learning through critical reflection. In *Innovative assessment in higher education* (eds. C. Bryan & K. Clegg), pp. 92-99. Routledge, London.

Rourke, L., Anderson, T., Garrison, D. R. & Archer, W. (2001) Methodological Issues in the Content Analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education* ***12***, 8-22.

Schellens, T. & Valcke, M. (2000) Re-engineering conventional university education: Implications for students' learning styles. *Distance Education* ***21***, 361-384.

Schellens, T. & Valcke, M. (2005) Collaborative learning in asynchronous discussion groups: What about the impact on cognitive processing? *Computers in Human Behavior* ***21***, 957-975.

Schellens, T., Van Keer, H. & Valcke, M. (2005) The impact of role assignment on knowledge construction in asynchronous discussion groups: A multilevel analysis. *Small Group Research* ***36***, 704-745.

Schellens, T., Van Keer, H., Valcke, M. & De Wever, B. (2007) Learning in asynchronous discussion groups: A multilevel approach to study the influence of student, group, and task characteristics. *Behaviour & Information Technology* ***26,*** 55-71.

Sluijsmans, D., Dochy, F. & Moerkerke, G. (1999) Creating a Learning Environment by Using Self-, Peer- and Co-Assessment. *Learning Environments Research* ***1***, 293-319.

Strijbos, J. W., Martens, R. L., Jochems, W., & Broers, N. J. (2004). The effect of functional roles on group efficiency: Using multilevel modelling and content analysis to investigate computer-supported collaboration in small groups. *Small Group Research* ***35***, 195-229.

Veerman, A., & Veldhuis-Diermanse, E. (2001). Collaborative learning through computer-mediated communication in academic education. In *Euro CSCL 2001*, pp. 625-632. Maastricht: McLuhan institute, University of Maastricht.

Weinberger, A., Reiserer, M., Ertl, B., Fischer, F. & Mandl, H. (2005) Facilitating collaborative knowledge construction in computer-mediated learning environments with cooperation scripts. In *Barriers and biases in computer-mediated knowledge communication* (eds. R. Bromme, F. W. Hesse & H. Spada), pp. 15-38. Kluwer, Boston.

Supplementary material

Translation of a discussion assignment from theme 4 (evaluation):

In this discussion assignment we focus on the assessment of the discussion groups. As you know (and as you can read in detail on the website) your contributions to the discussion groups are assessed by two instructors. This assessment is not only based on the number of contributions, but also (and more importantly) on the quality of your contributions. However, alternative approaches for assessment are possible as well. In this assignment we more specifically focus on the possibilities peer assessment. Your task is twofold:

- Discuss in group to which extent peer assessment could be an interesting supplement to the current assessment procedures of your online discussion assignments. Do not forget to debate the shifting insights concerning knowledge and learning goals. Make the advantages and disadvantages clear and focus on theoretical and/or practical drawbacks. Discuss the limiting conditions and focus on the feasibility of this type of assessment of the discussion groups.

- Develop – based on your discussion – a concrete peer assessment approach that could be used to assess the discussions. Situate this approach at the different dimensions [aggregation level, function, assessor, link learning goals, moment in time, assessment technique]. Focus more specifically on the function that you would assign to your peer assessment approach and link the assessment approach to the goals of the discussion groups explicitly.

Translation of a discussion transcript (selection from theme 4: evaluation)

Thomas (Message 10)

Peer assessment

Dear discussion group members, I would like to discuss the part of the assignment concerning peer assessment. Please complete.

Interesting supplement: Peer assessment could be an interesting supplement to the existing way of assessment because it is based on judgements of the ones that were actively involved in the discussions. The peers could point out which contributions are constructive and which 'beat about the bush', are lacking depth, involved sticking points, … It is good that the evaluation is not completely external. There is less discussion possible about the results.

Advantages: I discussed the advantages somewhat in the previous point.

Disadvantages (theoretical and practical drawbacks):
- I think that you need rather a lot of time to assess each person.
- It is possible that the one assessing does not know the theory well and this could lead to an unreliable assessment.

Limiting conditions: The peers should have enough knowledge on assessing and the matter that should be assessed. A computer and an internet connection are necessary.

► **Charlotte** (Message 10.1)

Additions peer-assessment

- Advantages
  - By constructing "group criteria" together you get **collaborative learning** (advantages of this: see constructivism)

  - The **own goals** are formulated explicitly and are discussed afterwards: the learners control the process. The responsibility is in our hands and thus we can direct the discussion to our own goals…
  - By collaborating in this way we could get a view on our **metacognitive knowledge** by which we could get a good look at ourselves.
  - Our **competences** are exposed by the others and by ourselves.

- Disadvantages
  - We need **marks** at the end of the semester: on what will they be based? On our goals or on the goals of the instructors?
  - The **time aspect** that Thomas wrote about is indeed a disadvantage. But is this not a part of the learning process? (Is this an advantage or a disadvantage?)
  - Is there a general way to **assess**?

- Limiting conditions:
  - In my opinion, there should be **trust** and **respect** in our discussion group.
  - The discussion may not be a personal attack, but an addition that is directed at '**unconditional acceptance**' and constructive feedback on everybody's learning process.
  - The activity of assessing should be taken **seriously** by every member of the discussion group. In this way you can formalize this kind of 'group assessment' and the firstly discussed disadvantage could be countered.

► **Ellen** (Message 10.2)

Addition peer-assessment

Advantages:

- You have a say, you are a part of the whole, in other words: you are **participating**.
- You get insight in your own **communicative skills**. You learn to formulate what you think. In my opinion, peer assessment has a surplus value with respect to 'learning to assess', which is a **learning activity** an sich in my opinion.

Disadvantages

- There is no **context** at all! In my opinion, this is the largest disadvantage.
- You have to interpret everything that is said or just not said. In my opinion, this is a disadvantage because in this way **objectivity** can not always be guaranteed.

Limiting conditions

- As I mentioned above, **objectivity** should be pursued.
- 'Peers' means a group of people that comply with equal values, norms, and visions. This group is formed on a **voluntary basis**. In my opinion, this is an important condition, but this is not the case with us.
- **Equivalence**. Both *formal as informal* no one should have a higher position. This is almost impossible to pursue, but I still think this should be taken into account.
- There should be **safety**. Everybody should be able and should dare to say what he/she wants.
- Everybody should **agree with** this way of assessment. This is related to the previous point.
- This form of assessment is usually used as a basis for a **formative assessment**. Thus, a summative, predictive, or selective evaluation may not be the basic assumption.
- To me the obvious means of performing peer assessment is not through a means of communication **but face-to-face**.

I indeed put a lot of limiting conditions and I link more disadvantages then advantages to this way of assessment in view of assessing the discussion groups. I am rather in favour of an assessment centre approach. The final assessment is not dependent on one person who is responsible. It is a combination of sub assessments.

One more question: Is it intended that peer assessment is done verbally, or written? I assume the former, but I have the feeling that in the course reader it is seen as a written process. I can not determine it very well. Can you?

►► **Dieter** (Message 10.2.1)

context

The context that is missing. What do you mean by this?

►►► **Charlotte** (Message 10.2.1.1)

There ìs context

Ellen, it is also not clear to me which context you are referring to. Because of the fact that everybody is situated in the same *'context'* of the discussion groups, I think it the surplus value is that we can put forward the same learning goals (as I said earlier).

We perfectly know which contributions to **our** discussion are interesting, superfluous, unfocused, a surplus value, … In this way, **we** can provide relevant critique with relation to the *content*, the way of putting it, … which I really see as an advantage.

►►►► **Ellen** (Message 10.2.1.1.1)

context

The context I am rather referring to is the personal background of everyone. I think that this may also fall under the denominator "context". Maybe everybody has a personal reason why he/she posts a lot of contributions or otherwise. You can not say or write this (for example within these discussion groups) en thus nobody is taking this into account, which is rather logical.

Sorry for the lack of clarity

►►►► **Aileen** (Message 10.2.1.1.2)

reaction to charlotte

I think it is dangerous to think that we know "perfectly" what is interesting and superfluous in our discussion. Do not forget that we work on this as an individual. What one believes to be interesting can be superfluous for somebody else. In my opinion, peer group assessment looks like a too subjective way of assessment. You can never guarantee that the assessment in a peer group is fair and honest.

I thus agree with Ellen that we should take our 'own' context into account. Everyone of us has had a preliminary training and we all went through different experiences. It is thus naïve to think that we are assessing all reactions in the same way. I think Peer group assessment that only takes place in the discussion groups is insufficient as assessment means.

Coding scheme and coding decisions

Gunawardena *et al.* (1997) reason that knowledge construction evolves through a series of levels. More specifically, five levels with subcomponents are distinguished. The first level is sharing and comparing of information, which comprises (1A) observations and opinions, (1B) statements of agreement, (1C) examples, (1D) clarifications, and (1E) identifications of problems. This is followed by level 2: the discovery and exploration of dissonance or inconsistency among ideas, concepts, or statements. This level comprises (2A) identification of disagreement, (2B) clarification of disagreement by asking and answering questions, and (3C) restating positions. The third level is negotiation of meaning and/or co-construction of knowledge, which includes (3A) negotiation or clarification of the meaning of terms, (3B) negotiation of the weights of arguments, (3C) identifications of areas of agreement or overlap among conflicting concepts, (3D) proposing and negotiation of new statements embodying compromise, and (3E) integrating metaphors or analogies. The fourth level is characterised by testing and modification of proposed synthesis or co-construction. These co-constructed statements are tested against (4A) shared "received facts", (4B) existing

cognitive schema, (4C) experiences, (4D) formal collected data, and (4E) literature. The fifth and final level refers to statements of agreement and application of newly-constructed meaning, and encompasses (5A) summarising agreements, (5B) applications of new knowledge, and (5C) metacognitive statements revealing new knowledge construction.

When applying this coding scheme, we made a three decisions based on the detailed coding scheme of Gunawardena *et al.* 1997 (p. 414). First, if students restate a position and bring in new arguments to support this position, this was coded as level 2 even though there was not always dissonance or inconsistency. This is in line with Gunawardena's *et al.* (1997) code 2C: "restating the participant's position, and possibly advancing arguments or considerations in its support by references to the participant's experience, literature, formal data collected, or proposal of relevant metaphor or analogy to illustrate point of view" (p. 414).

Second, if students refer to the course reader while testing and modifying proposed synthesis, this was coded as 4E: "testing against contradictory testimony in the literature" (Gunawardena *et al.* 1997, p. 414).

Third, we had to make decisions about summaries, since they could be used in different ways. When opinions or ideas were summarized in order to restate a position, the postings were coded as 2C: "restating the participant's position, and possibly advancing arguments or considerations in its support by references to the participant's experience, literature, formal data collected, or proposal of relevant metaphor or analogy to illustrate point of view" (Gunawardena *et al.* 1997, p. 414). If summaries served to negotiate about meanings, the postings were coded 3A ("Negotiation or clarification of the meaning of terms", Gunawardena *et al.* 1997, p. 414). If they served to identify areas of agreement or overlap among conflicting concepts they were coded 3C ("Identification of areas of agreement or overlap among conflicting concepts", Gunawardena *et al.* 1997, p. 414). Only when summaries reflect a higher level of consciousness and focus on the overall discussion process while summarizing agreements, they were coded 5A ("Summarization of agreement(s)", Gunawardena *et al.* 1997, p. 414).

These three decisions were based on the fact that the main intention of the posting should be reflected. During the discussion process, a lot of information is shared (level 1), but if the main purpose is to negotiate and convince someone of your case, then this should be coded as negotiation of meaning and/or co-construction of knowledge (level 3).

Overview of the different codes assigned to the messages

Thomas (Message 10)

1 Sharing / comparing of information

(1A A statement of observation or opinion)

► **Charlotte** (Message 10.1)

1 Sharing / comparing of information

(1A A statement of observation or opinion)

► **Ellen** (Message 10.2)

1 Sharing / comparing of information

(1A A statement of observation or opinion)

►► **Dieter** (Message 10.2.1)

1 Sharing / comparing of information

(1D Asking and answering questions to clarify details of statements)

►►► **Charlotte** (Message 10.2.1.1)

2 The discovery and exploration of dissonance or inconsistency among ideas, concepts or statements.

(2A Identifying and stating areas of disagreement)

►►►► **Ellen** (Message 10.2.1.1.1)

3 Negotiation of meaning / co-construction of knowledge

(3A Negotiation or clarification of the meaning of terms)

►►►► **Aileen** (Message 10.2.1.1.2)

4 Testing and modification of proposed synthesis or co-construction

(4C Testing against personal experience)

Statistical output

Statistical output research question 1: Students' ability to evaluate their own social knowledge construction processes

In order to explore whether students are able to assess their own level of knowledge construction through social negotiation, we focus on the match between students' self-assessment and the content analysis of their messages. Following Longhurst and Norton (1997) a convergence measure was computed per discussion theme by calculating the difference between the self-reported occurrence of utterances reflecting each level of knowledge construction (LKC_{SA}) and the observed occurrence of messages for each level of knowledge construction (LKC_{OBS}) as coded by the coders during the content analysis.

For each level of knowledge construction, a difference score was calculated (LKC_{DIF} 1 to 5). Negative difference scores indicate that students underestimate their level of knowledge construction, while positive scores point at overestimation. For each LKC_{DIF}, a three-level model was set up, in which the four successive discussion themes and self-assessment assignments (level 1) were nested within students (level 2), who were grouped themselves in discussion groups (level 3). First, random intercept null models were estimated. In a null model, the total variance of students' LKC_{DIF} is decomposed into between-group, between-students, and between-theme variance. Next, compound symmetry models were estimated for each LKC_{DIF}. These are random intercept models with no explanatory variables except for the measurement occasions (Snijders & Bosker 1999). They allow us to study the differences between the successive themes, by contrasting theme 2, theme 3, and theme 4 with the reference category (theme 1).

The random intercepts null models (null) and the compound symmetry models (CSM) for all five levels of knowledge construction can be found in appendix A. The null models indicate that variance in the difference score between discussion groups is low (0 % - 6 %), the variance between students within groups is medium (20 % - 30 %, except for LKC_{DIF} 4: 12 %), and the variance between themes within students is high (64 % or more). Furthermore, they indicate that students underestimate themselves at the first level of knowledge construction (LKC_{DIF} 1 = -1.103, *SE* = 0.128) and overestimate themselves at the four subsequent levels (LKC_{DIF} 2 = 1.314, *SE* = 0.086; LKC_{DIF} 3 = 1.344, *SE* = 0.101; LKC_{DIF} 4 = 2.280, *SE* = 0.084; LKC_{DIF} 5 = 1.714, *SE* = 0.067).

Figure 1 gives an overview of the difference scores with regard to the different levels of knowledge construction throughout the four themes (see also CSM models in the supplementary material). Dotted lines represent unsignificant differences, whereas full lines represent significant differences between consecutive discussion themes. Figure 6.1 clearly indicates that students underestimate themselves concerning the occurrence of the first level of knowledge construction reflected in their contributions and overrate themselves with respect to the following levels (2 till 5) in all themes. For the second level, students overestimate themselves significantly more in the second

theme (compared to the first theme). With regard to the third level, no significant differences were encountered between the four themes. Concerning the fourth level, a significant decrease in overestimation is noticed in theme 3 compared to theme 2 and a significant increase in overestimation in theme 4 compared to theme 3. Finally, in the fifth level a significant decrease in overestimation is observed between theme 2 and 1 and between theme 3 and 2.

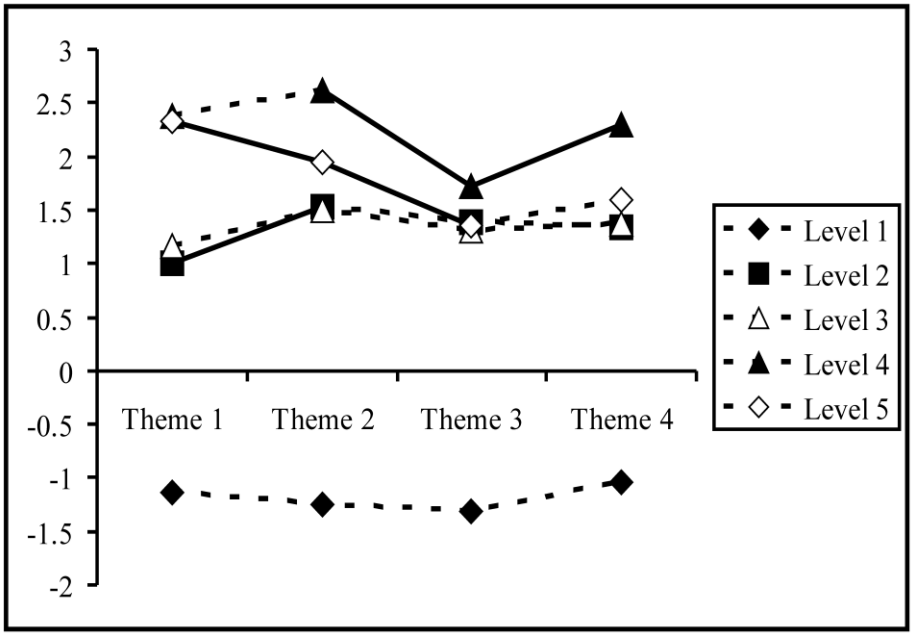

Figure 1. Graphical representation of the convergence measures (LKC_{DIF}) for each level of knowledge construction and each theme.

Statistical output research question 2:
Impact of the research conditions on levels of knowledge construction reflected in the online discussions

The second research question focuses on the impact of the three different research conditions on the levels of knowledge construction reflected in student messages (LKC_{OBS}). For this research question, a four-level model was estimated, with messages (level 1) hierarchically nested within measurement occasions (level 2) that are clustered within students (level 3) who are in turn assigned to groups (level 4). The analysis models were built following a stepwise procedure. Comparable to the first research question, a random intercept null model and a compound symmetry model were estimated first. Next, additional analyses were performed in which the different research conditions were included as predictors to the model. All models are presented in Appendix B.

The random part of the four level null model (model 0) for LKC_{OBS} shows that the variances on group, theme, and messages level are significantly different from zero: 4.89% of the total variance in LKC_{OBS} in students' messages is situated at the group level ($\chi^2 = 8.129$, $df = 1$, $p = .004$), 5.76% is situated at the theme level (measurements occasions) ($\chi^2 = 29.501$, $df = 1$, $p < .001$), and 89.35% of the variance arises from differences between messages within measurement occasions ($\chi^2 = 2060.958$, $df = 1$, $p < .001$). No part of the total variance can be assigned to the level of the individual students.

Next, a compound symmetry model (model 1) is compared with the null model, using the difference in deviance of both models as a test statistic having a chi-squared distribution with the difference in number of parameters as degrees of freedom (Snijders & Bosker 1999). The compound symmetry model achieves a better fit than the null model ($\chi^2 = 146.410$, $df = 3$, $p < .001$). Compared to theme 1, the LKC_{OBS} in theme 4 is not significantly different ($\chi^2 = 1.265$, $df = 1$, $p = .261$). However, messages in theme 2 reflect a significantly lower LKC_{OBS} ($\chi^2 = 13.188$, $df = 1$, $p < .001$), while messages in theme 3 reflect a significantly higher LKC_{OBS} ($\chi^2 = 78.783$, $df = 1$, $p < .001$).

In model 2 the differences between the three conditions across all themes are revealed by adding the explanatory variable 'research condition' to the fixed part of the model. This categorical variable is represented by two dummies: 'condition 2' refers to the role/no-role condition in which groups were assigned roles in theme 1 and 2 and 'condition 3' refers to the role/no-role+SA condition with similar role assignment in theme 1 and 2 and with the additional support of reflection through self-assessment. Both conditions are contrasted against the reference category representing the no-role/role condition (condition 1) in which role assignment was introduced in theme 3 and 4.

Model 2 has a significantly better fit ($\chi^2 = 18.000$, $df = 2$, $p < .001$) and indicates that messages in both condition 2 and 3 reflect significantly higher LKC_{OBS} compared to messages in condition 1 ($\chi^2 = 27.521$, $df = 1$, $p < .001$. and $\chi^2 = 14.463$, $df = 1$, $p < .001$ respectively). No significant difference between condition 2 and 3 was revealed ($\chi^2 = 2.290$, $df = 1$, $p = .130$).

In model 3 the difference between the conditions is explored more deeply by taking the interaction effects between the conditions and the themes into account. In this respect the differential progress in LKC_{OBS} in the different research conditions is studied. This model has a significantly better fit ($\chi^2 = 59.060$, $df = 6$, $p < .001$). The difference between the three research conditions is depicted in Figure 6.2. The trend indicating that students' contributions in general reflect higher LKC_{OBS} in conditions 2 and 3 compared to condition 1 (as revealed by model 2) is significant for the first theme: the LKC_{OBS} is significantly higher in both condition 2 (role/no-role) and condition 3 (role/no-role+SA) compared to the first condition (no-role/role) (respectively $\chi^2 = 11.725$, $df = 1$, $p = .001$ and $\chi^2 = 5.767$, $df = 1$, $p = .016$). No significant differences were found between condition 2 and 3 in theme 1 ($\chi^2 = 1.128$, $df = 1$, $p = .228$).

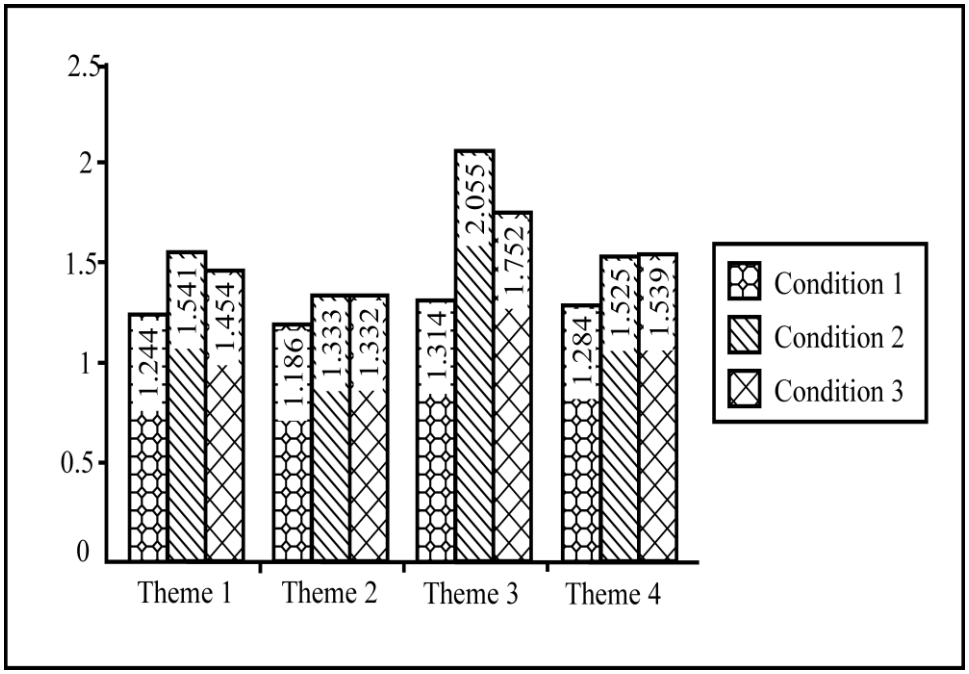

Figure 2. Graphical overview of the mean LKC_{OBS} per condition and per theme (based on model 3 in Appendix B).

The trend of higher LKC_{OBS} in the last two research conditions is also significant in the third theme: the LKC_{OBS} is significantly higher in condition 2 (role/no-role) and condition 3 (role/no-role+SA) compared to the first condition (no-role/role) (respectively $\chi^2 = 29.824$, $df = 1$, $p < .001$ and $\chi^2 = 7.853$, $df = 1$, $p = .005$). In addition, the LKC_{OBS} is significantly higher in condition 2 compared to condition 3 ($\chi^2 = 7.954$, $df = 1$, $p = .005$). The differences between the conditions are not significant for themes 2 and 4. As to the differences between role and no role based discussions, the results of model 3 indicate that for the initial discussion theme the discussion groups with role assignment reach higher levels of knowledge construction, whereas the opposite is true for theme 3. In condition 2 and 3 no role assignment was present in this third discussion theme and yet these discussion groups reach higher levels of knowledge construction.

Appendix A. Parameter estimates for the null models and compound symmetry models for the convergence measures for each level of knowledge construction (LKC)

Parameter	Null (LKC 1)	Null (LKC 2)	Null (LKC 3)	Null (LKC 4)	Null (LKC 5)
Fixed					
Intercept	-1.103 (0.128)	1.314 (0.086)	1.344 (0.101)	2.280 (0.084)	1.714 (0.067)
Random					
Level 3 – group					
σ^2v0	0.122 (0.101) 5.67 %	0.013 (0.044) 0.87 %	0.076 (0.062) 5.38 %	0.053 (0.043) 5.06 %	0.000 (0.000) 0 %
Level 2 – student					
σ^2u0	0.644 (0.187) 29.91 % ***	0.370 (0.130) 24.82 % **	0.303 (0.115) 21.44 % **	0.127 (0.078) 12.12 %	0.252 (0.085) 24.92 % **
Level 1 – theme					
σ^2e0	1.387 (0.154) 64.42 % ***	1.108 (0.122) 74.31% ***	1.034 (0.113) 73.18 % ***	0.868 (0.093) 82.82 % ***	0.759 (0.083) 75.07 % ***
Model fit					
Deviance	1080.109	969.748	955.492	869.910	855.241

Parameter	CSM (LKC 1)	CSM (LKC 2)	CSM (LKC 3)	CSM (LKC 4)	CSM (LKC 5)
Fixed					
Intercept	-1.134 (0.213)	1.010 (0.159)	1.156 (0.172)	2.371 (0.148)	2.332 (0.121)
Theme 2 (cognitivism)	-0.113 (0.233)	0.513 (0.205) *	0.345 (0.199)	0.239 (0.175)	-0.391 (0.150)
Theme 3 (constructivism)	-0.180 (0.226)	0.378 (0.199)	0.152 (0.196)	-0.654 (0.172) ***	-0.975 (0.148) ***
Theme 4 (evaluation)	0.091 (0.208)	0.330 (0.173)	0.207 (0.176)	-0.068 (0.152)	-0.741 (0.130) ***
Random					
Level 3 – group					
σ^2v0	0.160 (0.113)	0.013 (0.044)	0.078 (0.063)	0.063 (0.045)	0.002 (0.028)
Level 2 – student					
σ^2u0	0.642 (0.186) ***	0.402 (0.131) **	0.315 (0.116) **	0.135 (0.073)	0.345 (0.088) ***
Level 1 – theme					
σ^2e0	1.365 (0.152) ***	1.056 (0.117) ***	1.012 (0.111) ***	0.777 (0.084) ***	0.574 (0.065) ***
Model fit					
Deviance	1078.332	962.792	952.406	844.000	809.786

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

Appendix B. Multilevel parameter estimates of LKC_{OBS}

Parameter	Model 0 (null)	Model 1 (CSM)	Model 2	Model 3
Fixed				
Intercept	1.479 (0.044)	1.418 (0.050)	1.190 (0.057)	1.244 (0.065)
Theme 2 (cognitivism)		-0.132 (0.036)***	-0.131 (0.036)***	-0.058 (0.066)
Theme 3 (constructivism)		0.306 (0.034)***	0.308 (0.034)***	0.070 (0.061)
Theme 4 (evaluation)		0.040 (0.036)	0.043 (0.036)	0.040 (0.061)
Condition 2			0.376 (0.072)***	0.297 (0.087)***
Condition 3			0.272 (0.072)***	0.210 (0.088)*
Theme 2 * Condition 2				-0.150 (0.088)
Theme 2 * Condition 3				-0.064 (0.088)
Theme 3 * Condition 2				0.444 (0.081)***
Theme 3 * Condition 3				0.228 (0.081)**
Theme 4 * Condition 2				-0.056 (0.084)
Theme 4 * Condition 3				0.045 (0.084)
Random				
Level 4 – group				
$\sigma^2 f0$	0.034 (0.012)**	0.037 (0.013)**	0.013 (0.005)*	0.013 (0.005)*
Level 3 – student				
$\sigma^2 v0$	0.000 (0.000)	0.000 (0.004)	0.001 (0.004)	0.004 (0.004)
Level 2 – theme				
$\sigma^2 u0$	0.040 (0.007)***	0.017 (0.007)*	0.017 (0.007)*	0.006 (0.006)
Level 1 – message				
$\sigma^2 e0$	0.621 (0.014)***	0.618 (0.014)***	0.618 (0.014)***	0.617 (0.013)***
Model fit				
Deviance	11536.050	11389.740	11371.740	11312.680
χ^2		146.41	18	59.06
df		3	2	6
p		< .001	< .001	< .001
Reference		Model 0	Model 1	Model 2

Values between brackets are standard errors

* $p < .05$ ** $p < .01$ *** $p < .001$