Cooperative learning during math lessons in multi-ethnic elementary schools: counting on each other

Oortwijn, M.B.

CHAPTER 6

General conclusion and discussion

1. Introduction

In this thesis we carried out a study into the effectiveness of a CL math curriculum. Following Webb and Palincsar (1996), we distinguished four domains as regards the effectiveness of CL. We investigated the impact of teacher stimulation and two pupil background characteristics (ethnicity and prior knowledge) on math post-test scores and motivation for CL (Chapter 2), helping behavior (Chapter 3), and linguistic proficiency (Chapter 4) of 10-12 year old pupils from multi-ethnic classrooms. In Chapter 5, we studied how pupils' experience with CL influenced their popularity and perceived non-cooperativeness in general and how the CL experience affected ethnic differences in popularity and perceived non-cooperativeness in particular. We start this concluding chapter by linking the findings from the four different studies and extracting their key conclusions. Subsequently, we discuss the implications of these findings for multicultural elementary schools.

1.1 The hypotheses put to the test

The hypotheses that were investigated are depicted in Figure 2, Chapter 1. The first set of relations we investigated concerns the relationship between teacher stimulation and the pupils' math post-test scores and motivation for CL. In particular, we examined how these relationships were affected by pupils' ethnicity and their prior knowledge (both mathematically and linguistically). We hypothesized that teacher stimulation (experimental condition) raises the pupils' math post-test scores more when compared to the control condition (where pupils were left to fend for themselves), especially the math post-test scores of pupils with low prior math knowledge and of immigrant pupils. We found partial confirmation for this hypothesis: the math post-test scores of the pupils in the experimental condition were raised when compared with pupils' math post-test scores in the control condition. However, this only applied to the national pupils: the immigrant pupils did not have higher math gains in the experimental condition when compared with the control condition. This finding conflicts with Webb and Farivar's (1994) conclusion that immigrant pupils whose high-quality helping behavior is stimulated achieve higher learning gains, when compared with pupils whose high-quality helping behavior is not stimulated. In line with Gillies and Ashman (2000), we did demonstrate that teams with low prior math knowledge had higher math post-test scores in the experimental condition. We further hypothesized that pupils in the experimental condition are more motivated to cooperate than are

their counterparts in the control condition, especially immigrant pupils and the pupils with low prior math knowledge. The results were not clear-cut. We found that immigrant pupils with low prior math knowledge whose high-quality helping behavior was stimulated were more motivated to cooperate, which is in line with studies by Johnson and Johnson (2003) and Gillies and Ashman (1997). Unexpectedly, the immigrant pupils with high prior math knowledge reported higher motivation to cooperate in the control condition compared with their counterparts in the experimental condition. Post hoc analyses suggested that these pupils were linguistically less proficient, when compared with the national pupils with high prior math knowledge. This might have undermined their motivation to work in a CL setting where the pupils were required to participate actively in productive peer discussions.

In the second set of hypotheses we examined the quality of the verbal helping behavior of the pupils. We hypothesized that high-quality helping behavior (both solicited and unsolicited helping behavior) is positively related to learning gains. The data revealed that unsolicited helping behavior (tutoring behavior) was positively related to learning gains. This is in line with Topping (2005). The tutors were responsible for the positive effect of tutoring behavior on learning gains. We did not corroborate the findings of Webb and Mastergeorge (2003) who found a positive relationship between high-quality help and high-quality help application with learning gains. Remarkably, we found a negative relationship between asking for an explanation and learning gains. We further hypothesized that a possible advantage of national over immigrant pupils in the use of tutoring behavior and high-quality solicited verbal helping behavior is related to more limited linguistic proficiency of immigrant pupils. This explorative hypothesis was supported by the data: National pupils provided more tutor actions than immigrant pupils and this was associated with lower linguistic proficiency of immigrant pupils. Additionally, we hypothesized that pupils in the experimental condition use more high-quality solicited verbal helping behavior than the pupils in the control condition. We found no support for this hypothesis: There was no difference in pupils' use of high-quality helping behavior between the experimental condition and the control condition. This finding does not match with earlier studies that have suggested that stimulating pupils' high-quality helping behavior augments high-quality helping behavior (Fuchs, Fuchs, Kazdan & Allen, 1999). The last hypothesis in the second set stated that immigrant pupils and pupils with low prior math knowledge display more high-quality helping behavior in the experimental condition compared with the control condition. Furthermore, national pupils with low prior math knowledge were hypothesized to display more high-quality helping behavior in the experimental condition as compared to immigrant pupils with low prior math knowledge. The data did not support the hypothesis: Pupils with low prior math knowledge and immigrant pupils did not display more high-quality helping behavior in the experimental condition than in the control condition. Thus, we could not corroborate earlier

findings of Gillies and Ashman (1997; 2000) and Webb and Farivar (1994). We did find that the national pupils in the control condition made more use of low-quality helping behavior compared with the experimental condition.

The third set of hypotheses regarded the impact of teacher stimulation and pupils' ethnicity on their math-related linguistic proficiency. Firstly, we hypothesized that the pupils' math-related linguistic proficiency is positively related to their math post-test scores. The data revealed a negative relation between the math-related linguistic proficiency and the math post-test scores for immigrant pupils in the control condition, which is tentative support for the hypothesis. This finding is in line with other studies that have found that linguistic proficiency is related to academic development (e.g., Abedi & Lord, 2001; Cardelle-Elewar, 1992). Furthermore, we hypothesized that the pupils' math-related linguistic proficiency is higher in the experimental condition than in the control condition. The data supported this hypothesis, which is in line with earlier studies, like the study by Calderón et al. (1998). We also hypothesized that the math-related linguistic proficiency of immigrant pupils in the experimental condition is stimulated more than that of national pupils in the experimental condition. Indeed, we found that the peer talk of immigrant pupils in the experimental condition was characterized by a higher math-related linguistic proficiency as compared to the peer talk of immigrant pupils in the control condition. This finding extends the findings by Webb and Farivar's (1994) study, which showed that teachers who stimulated the pupils' use of high-quality helping behavior augmented the high-quality helping behavior and math gains of immigrant pupils.

The fourth set of hypotheses concerned the effect of CL experience on pupils' popularity and perceived non-cooperativeness. We hypothesized that the team members' popularity ratings increase and their perceived non-cooperativeness decreases in function of the time pupils spend working in their team. This hypothesis was confirmed: pupils perceived their fellow team members to be more popular and less non-cooperative at the end of the CL curriculum than at the beginning of the CL curriculum. This result confirms findings from earlier studies that have demonstrated that experience in CL strengthens the intragroup relations (Furrer & Skinner, 2003; Gillies & Ashman, 1997; Johnson & Johnson, 1994; Wright, Giammarino & Parad, 1986). Furthermore, we hypothesized that the popularity of immigrant pupils is increased and their perceived non-cooperativeness decreased after a CL experience. The data confirmed the first part of the hypothesis regarding the popularity and partly supported the second part regarding the perceived non-cooperativeness. In line with Levy, Kaplan, and Patrick (2004) and Prater, Bruhl, and Serna (1998) we found that the perceived non-cooperativeness of immigrant pupils decreased as a function of CL time. Unexpectedly, we found that the positive effect of CL time on the perceived non-cooperativeness was larger for national pupils than for immigrant pupils. Finally, we hypothesized that the popularity of the ethnically heterogeneous teams is stimulated and the

perceived non-cooperativeness is decreased by a CL experience. The data supported this hypothesis: We found that CL time decreased the perceived non-cooperativeness within ethnically heterogeneous teams and there was a trend for a positive effect of CL time on the popularity within ethnically heterogeneous teams. These findings extend the results from earlier studies that have demonstrated the positive effect of CL time on interethnic friendships (e.g., Slavin & Cooper, 1999).

2. General discussion

The following conclusions are drawn from the earlier mentioned findings: 1) the pupils' use of high-quality helping behavior needs to be supervised by the teacher to guarantee successful CL, 2) a CL experience strengthens pupils' social and academic performance, 3) there is a discrepancy between teachers' self-reported behavior and observed teacher stimulation, 4) ethnic background alone is not enough to predict social and academic performance in CL, and 5) heterogeneous grouping positively impacts on interethnic relations. Below, these conclusions are discussed in more detail.

2.1. Pupils' use of high-quality helping behavior needs to be supervised by the teacher to guarantee successful CL

Whereas Webb and Farivar (1994) found that stimulating pupils' use of high-quality helping behavior increased the performance of immigrant pupils, we found that the national pupils displayed higher learning gains as compared to the immigrant pupils. We found that immigrant pupils were less motivated to cooperate in the experimental condition than in the control condition. Why? In Chapter 3 we showed that immigrant pupils were less active during CL in terms of team conversations in the experimental condition. We also showed both in Chapters 2 and 3 that the linguistic performance of immigrant pupils was lower than that of the national pupils. Thus, limited linguistic proficiency may be argued to be the reason for the fact that immigrant pupils did not perform better in the experimental condition. Webb and Farivar (1994) did not report the linguistic proficiency of the immigrant pupils in their study. Perhaps this is because they did not score lower than the national pupils.

With respect to pupils' use of helping behavior, Chapter 2 and 3 showed that the teachers did not adequately stimulate pupils' high-quality helping behavior in the experimental condition. Thus, pupils did not learn to use more high-quality helping behavior, which may also be an explanation why we did not find a positive relation between pupils' high-quality helping behavior and their math post-test scores. The fact that teachers could not stimulate pupils' high-quality

helping behavior is reminiscent of Webb, Nemer, and Fall's (2006) study, which showed that pupils tend to mimic the behavior that the teacher displays. Nevertheless, the teachers in the experimental condition were able to decrease pupils' use of low-quality helping behavior (giving answers only) when compared with the control condition. Webb et al. (2006) provided an explanation for this finding. They suggested that teachers are more inclined to use corrective feedback in their teacher-student interactions during CL, because they are more familiar with this type of interaction. Consequently, they are effective in reducing this type of helping behavior in the pupil teams when asked to do so (as was the case in the experimental condition).

2.2. CL experiences strengthen pupils' social and academic performance

Two reasons are put forward here to explain the phenomenon that teacher stimulation did not result in more use of high-quality helping behavior by the pupils. The first is that the CL curriculum was too short for the pupils to integrate the high-quality helping behavior in their peer interactions. Other studies have demonstrated that CL experience of the pupils enhances academic performance (e.g., Gillies & Ashman, 1997). The majority of the pupils in the sample had no prior knowledge of CL skills. If the pupils in the experimental condition had been more experienced in CL, the difference in the quality of helping behavior between the pupils in the experimental condition and the pupils in the control condition might have been more striking. Nevertheless, the CL curriculum was long enough to bring about a positive change in the pupils' social status and skills. This suggests that the presence of high-quality helping behavior is not required to increase social status and skills. In addition to the relative short duration of the CL curriculum, another reason for the fact that the teachers could not boost the high-quality helping behavior of the pupils may be the minimal prior knowledge of CL skills and CL training of the participating teachers. Gillies (2003) suggested that most teachers lack the skills to implement CL successfully. It is likely that it was a lack of knowledge on the part of the teachers about how to stimulate pupils' use of high-quality helping behaviors that hindered the use of high-quality helping behavior by the pupils. The fact that the teachers brought down the frequency of pupils' use of low-quality helping behavior (asking for answers, and giving outcomes only) indicates that teachers can influence pupils' helping behavior in a relatively short time. Teachers with more experience in teaching pupils how to use high-quality helping behavior might have induced an advantage in high-quality helping behavior in the experimental condition over the control condition. The assertion that the teachers lacked the necessary CL experience necessitates a qualification of the theoretical model as proposed in Chapter 1, Figure 2. In addition to the pupils' prior knowledge of CL skills, we contend that the prior knowledge of CL skills of the

teacher is of equal importance. To test this, in future studies a closer look at the teacher' prior knowledge of CL skills is warranted (see also Webb et al., 2006).

2.3. Discrepancy between teachers' self-reported behavior and the observed teacher stimulation

In Chapters 2, 3, and 4 we manipulated teacher stimulation during CL by creating an experimental and a control condition. To ascertain whether the experimental and the control condition actually differed with respect to teacher stimulation, we checked the treatment integrity. Teacher stimulation was assessed in two ways. Firstly, teachers were required to indicate on a series of checklists how much they felt they were stimulating pupils' high-quality helping behavior. Secondly, the degree to which the teachers encouraged pupils' high-quality helping behavior was judged by the researcher and an independent observer by coding a number of videotaped CL lessons. The treatment integrity check revealed an interesting fact. The teachers reported implementing more high-quality helping in the experimental condition during CL when compared with the control condition. On the other hand, the two observers judged the teachers as using more whole-class reflections after the group work in the experimental condition compared with the control condition. This finding supports the explanation that the teachers did not have the know-how to successfully stimulate pupils' use of high-quality helping behavior. This result reflects findings from earlier studies (Sharan, 1990; Vedder & Veendrick, 2003), which suggested that there is a difference between what teachers think they are capable of and what they are actually doing. The fact that the teachers were not observably stimulating pupils' high-quality helping behavior during group work coincides with the finding that there was no difference between the experimental and the control condition with respect to pupils' use of high-quality helping behavior.

2.4. Ethnicity alone is not enough to predict social and academic performance in CL

In this thesis we found evidence that prior math knowledge and linguistic proficiency are better predictors of math post-test scores and high-quality helping behavior than ethnicity. Regarding linguistic proficiency, Chapters 2 and 3 showed that the immigrant pupils were considerably less linguistically proficient than national pupils. We suggest that the limited linguistic proficiency of immigrant pupils inhibited their active participation in the team. This suggestion is based on the following two findings: 1) in Chapter 2 we showed that a lower motivation for CL of immigrant pupils with high math prior math knowledge in the experimental condition was associated with a more limited linguistic proficiency as compared to national pupils with high prior math knowledge, and 2) in Chapter 3 we found that the degree to which

immigrant pupils assumed the role of tutor depended at least in part on their linguistic proficiency.

In addition to linguistic proficiency, prior math knowledge was the second contributor to differences between national and immigrant pupils in math post-test scores and helping behavior. Chapter 2 showed that national teams with high prior math knowledge had higher math scores in the experimental condition than national teams with low prior math knowledge. On the other hand, the math scores of immigrant teams with high prior math knowledge did not differ from immigrant teams with low prior math knowledge in the experimental condition. Furthermore, in Chapter 3 we found that immigrant pupils with medium-to-high prior math knowledge gave less high-quality help and displayed less low-quality constructive activity than national pupils with medium-to-high prior math knowledge. We argued that this finding is related to the poorer linguistic proficiency of the immigrant pupils. The requirement in the experimental condition to use specific high-quality helping behavior in peer talk hindered the performance of the immigrant pupils with medium-to-high prior math knowledge and undermined their motivation for CL. In the control condition they were not restricted in any way and thus had a higher motivation for CL. The opposite applied to immigrant pupils with low prior math knowledge. They had higher scores and higher CL motivation in the experimental condition than in the control condition, although their linguistic proficiency was not higher than that of the immigrant pupils with medium-to-high prior math knowledge. It could be that the immigrant pupils with low prior math knowledge were simply not able to comprehend the assignment on their own, in contrast to the immigrant pupils with medium-to-high prior math knowledge.

2.5. Ethnic heterogeneous grouping positively affects interethnic relations

In Chapter 5 we showed that the perceived non-cooperativeness is decreased and the popularity of ethnically heterogeneous teams is increased after a CL experience. This finding confirms Pettigrew's assertion that interethnic bias is reduced after prolonged interethnic exposure only. We assessed interethnic bias with a questionnaire on social status and skills, not by measuring friendships. Other studies have revealed that a CL experience affects interethnic social status development in the same manner as it affects the development of interethnic friendships (e.g., Warring, Johnson, Maruyama & Johnson, 1985). Thus, prolonged contact of team members with different ethnic backgrounds facilitates the interethnic friendships and consequently lowers the interethnic bias in these teams.

3. Limitations and guidelines for future research

This thesis had a number of limitations that need to be taken into account in future studies. First of all, as stated before, the CL curriculum was probably too short to find differences in the application of high-quality helping behavior by both the teachers and the pupils. Practical constraints were responsible for the length of the CL curriculum. Most notably the implementation of a longer, more intensive investigation was prevented by the limited time that the teachers had at their disposal and the high workload they experienced, particularly in multi-ethnic schools (Tesser & Iedema, 2001). We concur with other researchers (Cohen, 1994; Webb et al., 2006) that effective CL is a long-term process: it needs to be implemented over an extended time to assess its full effectiveness. Nevertheless, the gains in the pupils' social skills and status as well as math and linguistic performance in the short term testify to its potential.

A second limitation regards the small sample size. Due to this, multi-level analyses were not performed as they would have provided inaccurate results. Therefore, we analyzed the findings both at the individual level and at the team level by aggregating the individual variables. In any case larger samples are advisable to appropriately analyze the effectiveness of group work. However, this is only half of the story: the other half concerns randomization procedures. In this thesis, randomization was implemented at the classroom level. We were of the opinion that the limited resources of the teachers (time, prior knowledge of CL skills) would make it difficult to let them randomly apply structured or unstructured CL to the different teams within their classroom. Randomization at the team level would probably have decreased the chance of having a biased sample.

A third limitation regards the term 'ethnicity'. A study by Kistner et al. (1993) asserted that research into interethnic bias in multicultural classrooms should not only focus on ethnic status in terms of majority or minority per se, but also required that attention should be paid to the relative majority / minority ratio in a specific classroom. We only measured ethnicity as a dichotomous variable. In future studies it would be more informative to use both pupils' ethnicity in the classroom and the society at large. Furthermore, a number of Dutch studies have revealed that there are a number of academically distinguishable ethnic groups in the Netherlands (Tesser & Iedema, 2001). Thus, it would be interesting to investigate the impact of CL on pupils with different ethnic backgrounds.

The fourth limitation regards the assessment of the teacher background characteristics. We randomly assigned the teachers to either the control or the experimental condition. After being assigned, a workshop was held to instruct the teachers how to carry out effective CL without regard to their teaching styles. We did not adequately control for the teaching style that teachers used prior to the CL curriculum. Earlier research has suggested that teaching style is

hard to change (Gill, Ashton, and Algina, 2004; Webb et al., 2006). The study of Webb et al. (2006) resembles the CL study carried out in this thesis. They introduced CL to teachers who had no prior knowledge of CL skills and instructed them to target the high-quality helping behavior of their pupils (10 to 12 year olds). Their results showed that after an intensive training, the peer interactions were characterized predominantly by low-quality helping behavior -instead of high-quality helping behavior. These researchers concluded that teachers need specific training in questioning students and subsequently how to use the students' feedback to guide their joint problem-solving process. This is in keeping with our findings: that teachers themselves need to learn how to ask and give high-quality feedback before teaching their pupils in the use of high-quality helping behavior. Thus, we recommend that in future studies more attention is paid to the teaching style and what aspects of the teaching style need to be revised to assure that the group work is best served.