The impact of the teachter's role and pupils' ethnicity and prior knowledge on pupils' performance and motivation to cooperate

Oortwijn, M.B.; Boekaerts, M.; Vedder, P.H.

Instr Sci (2008) 36:251–268
DOI 10.1007/s11251-007-9032-7

The impact of the teacher's role and pupils' ethnicity and prior knowledge on pupils' performance and motivation to cooperate

Michiel Oortwijn · Monique Boekaerts · Paul Vedder

Received: 1 September 2006 / Accepted: 7 June 2007 / Published online: 29 June 2007

Abstract How can the teacher bring about effective cooperative learning (CL) in multiethnic elementary classrooms? To answer this question we hypothesized that when the teacher stimulates pupils' helping behaviour (experimental group), this increases pupils' performance and CL motivation more than when the teacher lets pupils fend for themselves (control group). Subjects were 166 pupils from 10 schools. The results show that national pupils in the experimental group outperformed pupils in the control group and teams with low and medium prior knowledge performed better in the experimental group. Additionally, immigrant teams with high prior knowledge in the control group outperformed their low prior knowledge counterparts and had a higher CL motivation. Our results suggest that, next to the teacher's role, attention has to be paid to both the pupil background characteristics ethnicity and prior knowledge and the teacher's experience with CL.

Keywords Cooperative learning · Teacher's role · Prior knowledge · Ethnicity · Elementary schools

Introduction

Cooperative learning is an educational method that has received a great deal of attention in the last decades (Cohen 1994; DeVries and Slavin 1978; Gillies 2004; Salomon and Perkins 1998). Following Cohen (1994) we define cooperative learning (CL) as an educational setting in which pupils work together in a group small enough that everyone can participate on a collective task that has been clearly assigned, without direct and immediate supervision from the teacher. Ample research has revealed that CL can be effective in enhancing the educational development of students (for meta analyses see Qin et al. 1995; Rohrbeck et al. 2003). However, CL is not per se more effective than direct teaching

M. Oortwijn (✉) · M. Boekaerts · P. Vedder
Faculty of Social and Behavioral Sciences, Centre for the Study of Education and Instruction, Wassenaarseweg 52, Leiden 2333 AK, The Netherlands
e-mail: oortwijn@fsw.leidenuniv.nl

methods (see for instance Pollock et al. 2002). In this study, we attempt to answer the following question: what can teachers do to make CL effective? In the remainder of this introduction we delineate the theoretical background of CL, followed by the three independent variables that we studied (the teacher's role, pupils' prior knowledge, and ethnicity) and conclude with the design of the study and our hypotheses.

Theoretical background of CL: the sociocultural approach

The history of research into CL traces back to Vygotsky's sociocultural approach (Vygotsky 1978/1930), who argued that learning is socially rooted. In his view, interactions with other learners in the social learning context are essential for the learner's development. Whether or not these interactions result in fruitful cognitive development depends on the level of cognitive development of each of the learners in the social learning context and the presence of an expert supervisor. The distance between the developmental age of the most capable learner and the least capable learner should not be too great, but just large enough for the least capable learner to benefit from the most capable learner. Vygotsky called this the zone of proximal developmental (ZPD). Within the ZPD the more capable learners support, or scaffold, the development of the lesser capable learners. In addition to this, research has demonstrated that the development of both types of learners can be facilitated by an expert supervisor, like a teacher (e.g., Gillies and Ashman 1997).

The teacher's role during CL

In this study we define the supervisor as the teacher. A number of studies have highlighted the importance of the teacher's role for successful CL (e.g., Gillies and Ashman 1997, 2000; Webb and Farivar 1994). The teacher's role refers, broadly speaking, to the educational tools the teacher applies to stimulate pupils' performance. Studies by Webb and her colleagues (Webb and Farivar 1994; Webb et al. 1995) have shown that pupils' performance is boosted when teachers encourage pupils to use high quality helping behaviour—defined as helping behaviour that includes asking for, providing, and applying explanations. Additionally, Chinn et al. (2000) reported that both cooperation and performance are boosted when the teacher promotes high quality helping behaviour. Following these researchers, we define effective teacher behaviour during CL as the ability of teachers to stimulate the high quality verbal helping behaviour of both individual team members and teams as a whole.

Pupil background characteristics

In the Netherlands, a significant percentage of the immigrant children on elementary schools have an educational disadvantage (Gijsberts 2004; Tesser and Iedema 2001). This, combined with the observation that the number of multiethnic schools is on the rise (Gijsberts 2004), calls for an examination of educational methods which can decrease the educational disadvantage in the multiethnic schools. Earlier studies have revealed that CL programs in which the teacher stimulates pupils' helping behaviour can be effective in reducing the educational disadvantage of immigrant pupils (e.g., Slavin and Cooper 1999; Webb and Farivar 1994). However, it is obvious that being an immigrant is not equal to being educationally disadvantaged. In this study, we argue that the relationship between

the teacher's stimulation of pupils' helping behaviour and ethnicity is affected by the level of pupils' prior knowledge.

Prior knowledge

Cohen, in her 1994 review study concludes that pupils with low prior knowledge benefit the most from CL in groups that are heterogeneous in prior knowledge. Puustinen (1998) argues that pupils with low prior knowledge are less able to self-regulate their learning. Consequently, they need more support from the teacher to actively participate in CL. In accordance with this, Gillies and Ashman (2000) showed that when teachers encourage pupils to use high quality helping behaviour, this stimulates the math performance of pupils with low prior knowledge more than when pupils are left to fend for themselves. Additionally, they demonstrated that when the teacher stimulates pupils' high quality helping behaviour the motivation to cooperate of pupils with low prior knowledge is augmented.

Regarding pupils with high prior knowledge the benefits of CL seem to be less clear-cut. Cohen (1994) suggests that these pupils needs less stimulation by the teacher to cooperate effectively. In keeping with this, Mulryan (1994) carried out a descriptive study to assess pupils' attitudes towards and behaviour during CL in five classes where pupils worked in groups on a regular basis. Interviews with the pupils revealed that pupils with high prior knowledge were more aware of the need of helping behaviour than low prior knowledge ones.

Ethnicity

Ethnicity is a second background characteristic that has been found to interact with the effectiveness of the teacher's role during CL. There is evidence that immigrant pupils in ethnically heterogeneous classrooms perform better when the teacher encourages them to use high quality peer interactions (Klingner et al. 1998). Webb and Farivar (1994) carried out a study in which the quality of helping behaviour that the teacher provided was manipulated: pupils were either encouraged in their application of the helping behaviour (experimental group) or not encouraged (control group). Their sample consisted of pupils from multiethnic elementary schools, of whom most immigrant pupils had an educational setback. Their study revealed that immigrant pupils in the experimental group outperformed the immigrant pupils in the control group. Other studies have reported similar findings (e.g. Calderón et al. 1998). In addition, reviews by Slavin and Cooper (1999) and Webb and Palincsar (1996) revealed that if the teacher encourages pupils' cooperativeness during CL interracial prejudice is reduced and the quality of cooperation augmented (see also Cohen 1994).

Research design, aim of the study and research considerations

This study concerns a comparative study that investigates the effect of the teacher's stimulation of the pupils' high quality helping behaviour on their performance and motivation to cooperate. We investigate this by manipulating the teacher's role: in the experimental condition the teachers are required to stimulate pupils' high quality helping behaviour as much as possible and in the control condition the teachers are required to let pupils fend for themselves as much as possible. The aim of this study is to corroborate the findings of Webb and Farivar (1994) and Gillies and Ashman (2000) that the teachers' stimulation of pupils' high quality helping behaviour augments their math performance. In addition to these studies we study how the effectiveness of the teacher's role during CL

interacts with ethnicity and prior knowledge. The effectiveness of the teacher's role is operationalized here as pupils' gain in math performance and perceived CL motivation.

Since the teachers in this study were required to address not only the individual team members during CL, but also the team as a whole, analyses will be performed both at the individual and at the group level.

Hypotheses

The following hypotheses are investigated: (1) When the teacher stimulates pupils' use of high quality helping behaviour (experimental group), the math performance of pupils is raised more than when they are left to fend for themselves (control group), especially that of pupils with low prior knowledge and of immigrant pupils; (2) pupils in the experimental group are more motivated to cooperate than are their counterparts in the control group, especially pupils with low prior knowledge and immigrant pupils.

Since the CL curriculum is in math, prior knowledge is defined here as prior math ability.

Method

Sample

A CL math curriculum of nine lessons was carried out in 10 multiethnic elementary schools. Letters were sent to 200 schools, of which 10 responded positively in a subsequent telephone conversation. Ten classes from 10 schools participated. In eight of these classes both the teacher and pupils had little or no prior experience with CL. Teachers of two classes (one in the experimental group, and one in the control group) indicated implementing group work frequently, around 80% of the time. Classes were randomly assigned to the experimental or the control group. The total sample consisted of 172 children. Six pupils who did not complete the math exam were dropped from the data set. As illustrated in Table 1, 166 children remained (average age 135.7 months, $SD = 6.5$), of which 71 were placed in the control group (57.7% male, 42.3% female), and 95 were placed in the experimental group (44.5% male, 55.5% female). With respect to ethnicity, pupils were regarded as national when both parents were of Dutch origin, "mixed" when one parent was of Dutch origin and "immigrant" when both parents were of foreign origin.

Table 1 Sample characteristics

Group	*N*	Gender	Ethnicity
Control group	71	30 female	16 national
			8 mixed
		41 male	47 immigrant
Experimental group	95	52 female	40 national
			21 mixed
		43 male	34 immigrant
Total	166		

"mixed" = one parent is Dutch, "immigrant" = both parents are of foreign origin

In the control group, 16 pupils were national, eight had a mixed background and 47 were immigrant pupils. The experimental group consisted of 40 national pupils, 21 mixed, and 34 immigrant pupils. Because there were relatively few pupils categorized as mixed in both conditions, the ethnicity variable was recoded into a dichotomous variable. The mixed pupils were classified as national pupils. Dutch studies (e.g., Tesser and Iedema 2001) have shown that the school performance of mixed pupils resembles that of national pupils more than that of immigrant pupils.

Procedure

The CL curriculum was composed of three phases. The first phase consisted of a mini-workshop of about 2 h, in which the researcher explained to the teachers, irrespective of group, exactly what the CL curriculum was about. Next, the teachers were given the most essential group specific instructions. Subsequently, the teachers were asked to carefully read the lesson-to-lesson CL protocol and they were encouraged to ask clarifying questions prior to the first CL training lesson.

In the second phase, which was identical for both conditions (experimental and control), all teachers placed their pupils in teams that were homogeneous narrow-heterogeneous in math ability (high-middle, or low-middle). Then the teacher trained the children to cooperate effectively in two 1-h lessons. In lesson 1, general social CL rules were taught and practiced. These rules were: "everyone cooperates", "everyone listens to each other", "everyone shares their knowledge and opinions", and "checks whether everyone agrees" following Webb et al. (1995). In lesson 2, more specific CL rules were mentioned and practiced. Adapted from Webb and colleagues (Webb and Farivar 1994; Webb et al. 1995), these rules all dealt with giving and receiving help. With respect to asking for help, we distinguished; (a) ask precise questions, (b) continue asking in case of ambiguities, (c) think before asking a question, (d) ask for help on time. With respect to giving help, we distinguished; (a) fine-tuning of the level of guidance to the need for help that is requested, (b) giving a clear and precise answer, (c) giving the help receiver a chance to apply the help given, (d) continuing to ask if the question for help is unclear and (e) giving help when needed. All CL rules (both the general and the more specific CL rules) were written down on a poster, which was displayed in the classroom and was clearly visible to all children of all classes of both conditions. This poster remained there throughout the whole CL curriculum as a memory aid for the pupils.

In addition to the poster, another memory aid for the pupils was a short checklist which they were required to fill in during each lesson. It also served as a check for pupils to see for themselves which CL rules they used inadequately. These checklists asked for the level of application of the general social CL rules that were taught in lesson 1 of the CL training and the amount of help given and help received (lesson 2 of the CL training). These checklists were not used for analysis.

Phase three consisted of a CL math curriculum of nine 1-h lessons, covering 5 weeks. The teachers carried out two lessons per week.

Experimental group

Following Webb and Farivar (1994), the impact of the teacher's stimulation of pupils' high quality helping behaviour on pupils' math performance was investigated (see Table 2). Two groups were created: an experimental and a control group. In both conditions the

Table 2 Summary of the CL curriculum

		Duration	
		Control group	Experimental group
A. Teacher activity			
1. Appointment of chairmen by the teacher		No	Yes
2. Direct teaching episodes	During CL to restore order	Yes	Yes
	Evaluation of the group work	No	Yes
3. Providing CL feedback (circulating among the groups)	Repeat + explain general basic CL rules	Yes	Yes
	Verbally rewarding use of rules for help giving and receiving	No	Yes
		No	Yes
B. Task structure			
4. Group assignment checks?		Yes	Yes
5. Explicitly mention in the assignments of:			
a. The responsibility of the chairman as role divider		No	Yes
b. The need to share and discuss the solutions		No	Yes

teachers carried out a CL math curriculum. However, teachers in the control group were trained to do nothing to stimulate pupils' high quality helping behaviour. The teachers only managed the group work when team members; (a) talked too loudly (disturbed other groups), (b) did not listen to each other, (c) made fun of each other. In contrast, the teachers in the experimental group stimulated pupils' high quality helping behaviour as much as possible. The teachers received a CL protocol that was group specific: teachers in the experimental group had a different protocol than the teachers in the control group. However, in both protocols detailed descriptions of all the math assignments were provided. Additionally, the protocol contained detailed lesson-to-lesson instructions about how the teachers should apply the CL rules in their group. In this way, differences between the experimental and the control group were optimized, enabling a better assessment of the effect of the teacher's stimulation of pupils' high quality helping behaviour on math performance.

Task structure

The assignments dealt with surface, percentage, scale, estimation and fractions. All assignments were adapted for CL purposes from a regular math curriculum that employs realistic math problems with a narrative composition. The assignments were moderately structured, open-ended, narrative math assignments, consisting of three parts. First, team members had to individually work on a part of the math task. Second, they had to discuss their findings. Third, all team members were required to cooperate to solve the last part of the math task. The teachers were asked in the protocol to emphasize in their instructions to the pupils that the focus in the CL curriculum was on understanding the math tasks rather than completing them.

Reward structure

The pupils completed an individual exam at the end of the CL curriculum (i.e., the math post-test). During the curriculum the teachers in both conditions took in the worksheet of a

random chosen team member of every team at the end of each lesson. The teachers in the experimental condition were required to provide feedback on the worksheets of each group, by verbally praising each group on one or more aspects of their helping behaviour. The worksheets were only discussed in the classroom: they were not taken into account when pupils' scores on the math exam were calculated.

Instruments

To check the integrity of the manipulation we used a teacher checklist of helping behaviour and videotaped teacher–pupil interactions. The pupils' ability was tested with a math pre and post-test and a pupil questionnaire on the quality of CL.

Manipulation integrity check

The manipulation integrity check consisted of a teacher checklist of helping behaviour and videotaped teacher–pupil interactions. Regarding the teacher checklist of helping behaviour, teachers rated on a 4-point Likert-scale (1 = "very often" and 4 = "very little") the extent to which they had implemented a number of CL rules. The teachers completed the checklist at the end of every week. A principal component analysis with varimax rotation revealed a three-factor solution. The solution explained 71% of the variance. Dimension one (18 items, $\alpha = .97$) entailed statements on the degree to which the teacher taught pupils the *use of general social CL rules*. A sample item is "I teach the children not to interrupt each other". Dimension two (five items, $\alpha = .81$) concerned the degree to which the teacher stimulated the *pupils' use of high quality helping behaviour*. A sample item is "I teach the children to keep asking when someone asks an unclear question". The third dimension (four items, $\alpha = .84$) covered to what extent the teacher gave *feedback on the group process*. A sample item is "At the end of each lesson I discuss with each group what is going well and what should be improved". In total, each teacher completed five checklists. Next to this questionnaire, teachers were required to indicate whether they implemented CL during regular lessons and whether they made more use of CL for the regular program during the CL curriculum than before the CL curriculum.

The second part of the manipulation integrity check consisted of videotaped teacher–pupil interactions. All teachers were videotaped during two or more lessons to know whether the two conditions differed regarding the implementation of CL rules. All recordings were rated by two independent scorers, one of whom was double blind to the experimental manipulation. The coding scheme comprised 14 items. A principal component analysis with varimax rotation was applied. It revealed a two-factor solution, explaining 62% of the variance. All factor loadings were .50 or higher. The first dimension (six items, $\alpha = .71$) was about the presence of *feedback on the group process*. A sample item is "Does the teacher reflect on group performance in the previous lesson?" The second dimension (eight items, $\alpha = .86$) covered items that were about the *teacher's whereabouts and activities during group work*. A sample item is "Does the teacher encourage team members to ask each other clear questions?" The items were rated on a 3-points scale. The higher the score, the more the teacher was perceived to encourage pupils' high quality helping behaviour. Due to technical failure, recordings were available of eight teachers only. In all, 18 recordings of teacher–pupil interactions could be coded. The overall inter coder reliability was assessed on the basis of Cohen's κ, calculated on two recordings (approximately 10%) and was found to be satisfactory: $\kappa = .68$.

Prior math knowledge (math pre-test)

Scores from a curriculum independent math test (CITO; Janssen et al. 1996) were used to assess the baseline math performance of all pupils. Since the teachers did not provide us with the data needed for the calculation of the internal validity, we refer to earlier research which has demonstrated that the curriculum independent math test has a good reliability, $\alpha = .94$ (Evers et al. 2000). Because some schools only provided the standardized 3-point scores (low, medium, high), all CITO scores used in this study were transformed into this 3-point rating scale. A Pearson's correlation test showed that the pre-test significantly correlated with the post-test, $r = 0.86$, $p < .001$.

Math post-test

This is an exam (with possible scores ranging from 1 to 10) that consists of seven items covering the math domains that the children learned during the math curriculum. All teachers applied the same curriculum-dependent math exam after the CL curriculum. A reliability analysis of the data obtained in this study revealed that the internal consistency was satisfactory, $\alpha = .75$.

Questionnaire for pupils on the quality of cooperative learning (QCL)

This is a questionnaire adapted from Hijzen et al. (2006). Items of the original questionnaire, which was intended for pupils from secondary vocational education, were reformulated for elementary school age pupils. It consists of two dimensions: "CL instruction" and "CL motivation". The pupils filled in the dimension CL instruction only once: before the start of the CL curriculum, to check for differences in CL experience between conditions. The dimension CL motivation was filled in twice, namely prior to the CL math curriculum, but after the CL training (T1), and a second time after the math exam (T2).

In total, the QCL consists of 30 items. A 4-point Likert-scale (1 = "very often" and 4 = "very little") was chosen instead of 5 points to avoid pupils opting for the middle, neutral category. The dimension "CL instruction" is made up of three scales. The scale "learned CL rules" (six items, $\alpha = .72$) is about pupils' perception of CL rules learned from their teacher (eigenvalue = 2.6, explained variance = 37.7%). A sample item is "The teacher has taught us to listen to the other team members during group work". The scale "planning of CL" (nine items, $\alpha = .81$) covers pupils' opinion about the teacher's preparation for group work (eigenvalue = 3.56, explained variance = 39.5%). A sample item is "Before beginning, the teacher tells us what we have to learn from the task". The scale "activity of the teacher during CL" (seven items, $\alpha = .75$) is about how pupils perceive the role of the teacher during group work (eigenvalue = 2.8, explained variance = 40.5%). A sample item is "During group work, the teacher frequently asks how we are getting along with the task". The dimension "CL motivation" (eight items, $\alpha = .83$) is about pupils' motivation to cooperate (eigenvalue = 3.7, explained variance = 46.1%). A sample item is "I think it's more fun to work together than to work alone".

An overview of all experimental measures is presented in Table 3.

Results

We started the analyses by checking differences between conditions in teachers' and pupils' experiences with CL, differences in prior math knowledge, and assessing the

Table 3 Overview of the instruments used in this study

Implemented instruments	Purpose	Number and times of measurement
Videotaped teacher–pupils interactions	Manipulation check	Two recordings during two lessons of the CL curriculum
Teacher checklist of helping behaviour	Manipulation check	Five measurements At the end of every other lesson
Questionnaire on CL (QCL)	Measuring pupils' perceived quality of CL	Part A (learned CL rules, planning of CL, and activity of the teacher during CL): one measurement, prior to the CL curriculum Part B (CL motivation): two measurements, one prior to the CL curriculum and one after the curriculum
Pre-test math scores	To assess the math knowledge of pupils prior to the CL curriculum	One measurement Before the start of the CL curriculum
Post-test math scores	Exam covering the math domains of the CL math curriculum	One measurement After the CL curriculum

manipulation integrity. Then we proceeded with the analyses of the relation between the independent variables "group" (experimental group or control group), "ethnicity", and "prior math knowledge" with the dependent variable "math performance", both at the individual and the group level. Next, we examined whether there were any differences in pupils' CL motivation that were associated with the group they were in, their ethnicity, and their prior math knowledge, again both at the individual and the group level.

Experiences with CL of teachers and pupils

An independent samples *T*-test showed that the two groups did not differ with respect to teachers' self-reported implementation of CL during regular lessons. Also, although teachers in both the experimental and the control group indicated using CL more often for the regular program during the CL curriculum than prior to the CL curriculum, the two groups did not differ from each other.

There were also no differences at the start of the CL curriculum regarding pupils' perception of CL instruction between the two groups.

Manipulation integrity

To assess the manipulation integrity we used the teachers' checklist of helping behaviour and the video recordings of teacher–pupil interactions. The teachers in the experimental group reported instructing pupils significantly more in the use of high quality helping behaviour, $t(21) = -3.37$, $p < .005$, Cohen's $d = 1.48$, than the teachers in the control group. No differences were found on the dimensions "general social rules" and "feedback on the group process".

Analysis of videotaped teacher-behaviour during CL lessons revealed that, although the sample was small ($N = 18$), homogeneity of variance did not differ between the two groups, the kurtosis was normal and the data were not significantly skewed. Analysis of the coded lessons showed that teachers in the experimental group did provide more feedback on the group process than teachers in the control group, $t(16) = -1.78$, $p < .05$, Cohen's $d = .58$. No differences were found for the dimension "CL activities during group work".

Quantitative analyses: putting the hypotheses to the test

Hypothesis 1: math performance of pupils in the experimental group is raised more than of pupils in the control group, especially that of pupils with low prior knowledge and of immigrant pupils. A repeated measures test was performed. The independent variables were ethnicity, group (experimental or control group), and prior math knowledge (low, medium, high). Checks of the assumptions showed that normality, linearity, and homogeneity of variance were satisfactory. No univariate or multivariate outliers were found. No main effects were found (see Table 4). We did find a significant two-way interaction effect for "group" × ethnicity, $F(1,161) = 4.51$, $p < .04$, explaining 3% of the variance [$\eta^2 = .03$] (see Fig. 1). This means that the post-test math scores of immigrant pupils in the control group were significantly better than that of national pupils in the control group, $F(1,68) = 5.9$, $p < .02$, $\eta^2 = .08$. Also, post-test math performance of national pupils in the control group was significantly lower than of national pupils in the experimental group,

Table 4 Mean scores of the pupils on the math post-test

Group	Prior math knowledge	Mean individual scores (*SD*)	Mean team scores (*SD*)
Experimental group			
National	Low	3.74 (1.16)	4.35 (.40)
	Moderate	5.32 (1.67)	5.42 (.79)
	High	6.82 (1.51)	5.92 (1.18)
	All	5.59 (1.94)	5.46 (1.05)
Immigrant	Low	3.86 (1.69)	5.81 (.05)
	Moderate	5.03 (1.24)	4.64 (1.19)
	High	6.01 (1.63)	4.7*
	All	4.78 (1.72)	4.94 (1.05)
Control group			
National	Low	2.64 (1.23)	3.43 (.83)
	Moderate	3.53 (.63)	2.6*
	High	6.09 (1.62)	5.56 (1.03)
	All	4.37 (2.10)	4.52 (1.51)
Immigrant	Low	3.53 (1.46)	3.62 (1.56)
	Moderate	5.01 (2.27)	4.36 (1.07)
	High	7.02 (1.52)	7.66 (.75)
	All	5.31 (2.27)	5.21 (2.12)

*These cells consisted of only one group. Therefore, the *SD* could not be calculated. Removal of these single measurements from analysis did not alter the significant finding

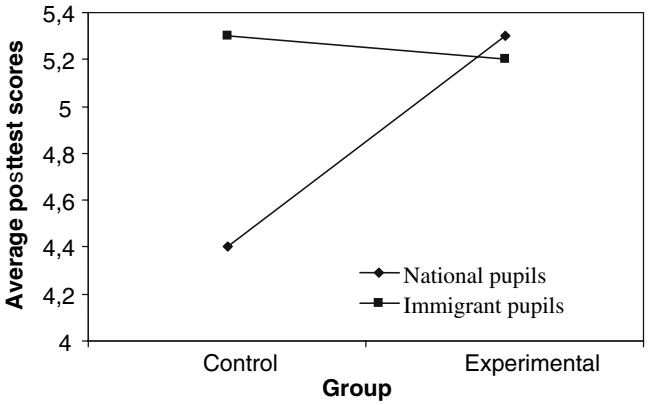

Fig. 1 Interaction "group" × "ethnicity" at the individual level

$F(1,82) = 8.02$, $p < .007$, explaining 9% of the variance [$\eta^2 = .09$]. Thus, national pupils did perform as we hypothesized, showing higher learning gains in the experimental group. In contrast to our hypothesis, immigrant pupils performed better in the control group. Furthermore, we could not demonstrate a positive effect of the stimulation of high quality helping behaviour on the performance of pupils with low prior math knowledge.

Analyses at the group level

Since pupils completed the math exam individually, the math scores of the individual pupils were used as unit of analysis. Nevertheless, as group work was the focus of research in this study, an explorative analysis of the math performance at the group level was also incorporated. However, due to the small sample size, the teacher's role could not be evaluated with a multilevel approach. Inspired by Gillies and Ashman (2000), Webb and Farivar (1994), and Saleh et al. (2005), analyses at the group level were executed by aggregating pupils' math scores from the pre-test as well as the post-test: scores and dividing these by the number of pupils in the teams. Regarding "ethnicity", a new variable was created (1 = majority of children have at least one Dutch parent, 2 = majority of pupils have immigrant parents). Also a new variable was created for "prior math knowledge" (1 = mean group pre-test math score is below average, 2 = mean group pre-test math score is on average, 3 = mean group pre-test math score is above average).

The group data ($n = 48$) were analysed with a repeated measures, in which "group" (experimental or control group), "ethnicity", and "average prior math knowledge" were the independent variables. No main effects were found. However, a significant two-way interaction effect was found for ethnicity with "group", $F(1,36) = 5.04$, $p < .04$ [$\eta^2 = .12$], indicating that teams with national pupils performed better when the teacher stimulated their use of high quality helping behaviour than when the teacher did not stimulate their use of high quality helping behaviour. Furthermore, a significant two-way interaction effect was found for "group" × "average prior math knowledge", $F(2,36) = 4.55$, $p < .02$, explaining 20% of the variance [$\eta^2 = .20$] (see Fig. 2). Teams with high prior math knowledge only performed better than teams with low prior math knowledge in the control group, $F(2,18) = 11.8$, $p < .005$. Because of the small sample, a Kruskal–Wallis test was carried out to cross-validate this finding, which yielded a similar result, $\chi^2 = 11.03$, *df* 2,

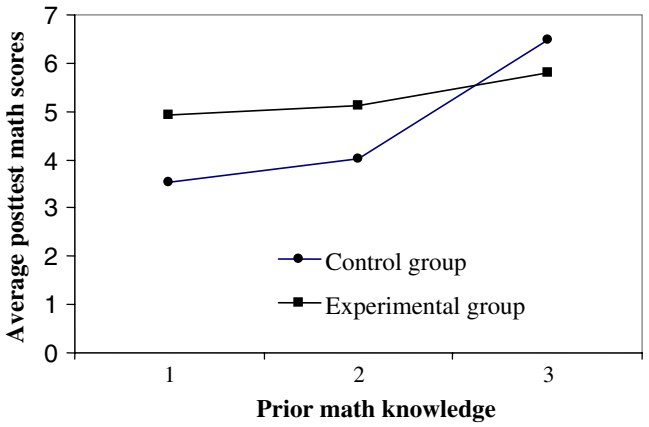

Fig. 2 Interaction "group" × "prior math knowledge" at the group level

$p < .005$. No differences between teams with high and low prior math knowledge were present in the experimental group. Furthermore, a Mann–Whitney test revealed that the stimulation of high quality helping behaviour was only positively related to post-test math performance of teams with low prior math knowledge, $Z = -2.30$, $p < .05$. For teams with medium prior math knowledge a trend was found, $Z = -78$, $p < .08$. Finally, a three-way interaction effect was found for "group" × "average prior math knowledge" × "ethnicity", $F(2,36) = 3.26$, $p = .05$, which explained 15% of the variance [$\eta^2 = .15$]. A Kruskal–Wallis revealed that teams consisting of national pupils with high prior math knowledge performed better in the control group than teams with national pupils with low prior math knowledge, $\chi^2 = 6.04$, df 2, $p < .05$. The same pattern emerged in the experimental group, $\chi^2 = 6.05$, df 2, $p < .05$. For teams with immigrant pupils, the picture was slightly different: there was a difference between immigrant teams with high and low prior math knowledge in the control group, $\chi^2 = 7.73$, df 2, $p < .03$, but not in the experimental group. It seems that teams with low prior math knowledge are generally better off in the experimental group, which is in line with our expectations. Unexpectedly, the immigrant teams with high prior math knowledge did not outperform the immigrant teams with low prior math knowledge in the experimental group.

Hypothesis 2: pupils in the experimental group are more motivated to cooperate than are their counterparts in the control group, especially pupils with low prior knowledge and immigrant pupils. In addition to six pupils who filled in only one questionnaire, one class unwittingly filled in the second questionnaire only. Therefore, the sample amassed 149 pupils: 67 in the control group (21 national, 46 immigrant) and 82 in the experimental group (58 national, 24 immigrant). To check for initial differences between the two groups, scores on the scales "learned CL rules", "planning of CL rules" and "activity of the teacher during CL" of the dimension "CL instruction" were compared between the two groups prior to the CL curriculum (T1) using a MANOVA (see Table 5). No differences on these scales between the two groups were found prior to the CL curriculum.

The effect of the CL curriculum on pupils' self-reported CL motivation was analysed at T1 and after the curriculum (T2) with a repeated measures test in which the independent variables were "group", "ethnicity" and "prior math knowledge". No significant main effects were found (see Table 5). However, the analysis did reveal a significant two-way interaction effect, Wilks' $F(2,130) = 3.20$, $p < .05$, which explained 5% of the variance

Table 5 Comparison of the control and experimental group with respect to pupils' scores on the QCL at T1 and T2

Dimension	Scale	T1		T2	
		Control (*SD*)	Experimental (*SD*)	Control (*SD*)	Experimental (*SD*)
1 (CL instruction)	Learned CL rules[a]	3.54 (.36)	3.53 (.36)		
	Planning of CL[a]	3.31 (.47)	3.32 (.61)		
	Activity of the teacher during CL[a]	3.34 (.51)	3.43 (.50)		
2 (Use of CL skills)	Application of general CL rules	3.66 (.44)	3.60 (.38)	3.69 (.42)	3.56 (.51)
	Application of helping rules	3.59 (.35)	3.51 (.45)	3.64 (.40)	3.48 (.48)
3 (CL motivation)	CL motivation	3.50 (.52)	3.34 (.59)	3.46 (.56)	3.19 (.66)

Higher mean scores indicate a higher perceived quality of CL

[a] Only filled in by pupils prior to the CL curriculum to check for initial differences

[$\eta^2 = .05$]. That is, pupils with low prior math knowledge in the experimental group were more motivated to cooperate at the end of the CL curriculum than pupils with high prior math knowledge in the experimental group. Further analysis showed that the effect was related to ethnicity, $F(2,59) = 3.78$, $p < .05$ [$\eta^2 = .11$]. That is, whereas immigrant pupils with high prior math knowledge were more motivated to cooperate in the control group, immigrant pupils with low prior math knowledge were more motivated to cooperate in the experimental group.

The fact that immigrant pupils with high prior math knowledge were more motivated to cooperate when they received no stimulation of their high quality helping behaviour resembles the analyses of pupils' math performance, which showed that immigrant pupils had higher math scores in the control group as compared to the experimental group. At the same time, these findings oppose our expectations. To understand why, we explored these findings further.

In the present study, the main difference between the two groups was the stimulation by the teacher of pupils' high quality helping behaviour. Intensive peer interaction presupposes a reasonable command of the language. Immigrant pupils are known to perform less well especially on tests of linguistic ability (Tesser and Iedema 2001). We analysed whether there were differences in linguistic proficiency between national and immigrant pupils, split by prior math knowledge (low, medium, high). "Linguistic proficiency" was measured prior to the CL curriculum with the scale "vocabulary" of the CITO, a national testing service in the Netherlands (Janssen et al. 1996). The analysis revealed an effect for the high prior math knowledge pupils only, $t(47) = 4.50$, $p < .001$: immigrant pupils with high prior math knowledge had a lower linguistic proficiency than the national ones. Next we checked whether the lower linguistic proficiency of the immigrant pupils with high prior math knowledge was related to their CL motivation. The earlier mentioned repeated measures analysis was repeated, but now with linguistic proficiency added as a covariate. The analysis showed that with the addition of "linguistic proficiency", the significant effect disappeared. Thus, immigrant pupils with high prior math knowledge may have had a lower CL motivation in the experimental group because of their more limited linguistic proficiency.

We also added linguistic proficiency as a covariate to the analyses of the relationship of group (experimental or control) with post-test math scores to explore whether this might

explain the unexpected findings. With the addition of this covariate the relation of the interaction of group and ethnicity with math performance was weakened. Furthermore, the linguistic proficiency of immigrant pupils with high prior math knowledge was lower than that of national pupils with high prior math knowledge in both the control and the experimental group, respectively, $Z(29) = -2.87$, $p < .005$, and $Z(20) = -2.38$, $p < .02$.

Analyses at the group level

Analyses to examine pupils' perception of CL at the group level revealed no significant effects. Therefore our prediction that teams in the experimental group would be more motivated to cooperate could not be confirmed.

Qualitative analysis of the effect of prior knowledge with ethnicity on peer interactions

We examined the interaction effect of ethnicity with prior knowledge more closely, by exploring pupils' use of peer interactions, more specifically their helping behaviour. We selected the interaction excerpts of two immigrant teams that differed in their use of helping behaviour. One team (team A) had a low average score in prior math knowledge, the other team (team B) a high average score in prior math knowledge. We explored to what extent the average level of prior math knowledge influenced the use of helping behaviour in these two teams. Both these teams were part of the experimental condition. In the excerpts, both teams were working on an identical assignment.

Generally, the peer interactions of team A were characterized by more high quality helping behaviour (for instance, giving an explanation—group mean 4.9—and applying the received help—group mean 3.4). They had an average CL motivation score of 3.8.

Pupil M: I don't understand 1B. What do I need to do here?
Pupil S.: Do you know what surface is, M.?
Pupil M: Yes, this side and this side (points out on his worksheet). For instance in this classroom the ceiling and sides (points at these)
Pupil S: But you're telling me nothing.
Pupil M: (shrugs shoulders)
Pupil S: Look, surface is length times width. For instance, you have a length of 3 meters and a width of 2 meters (draws a rectangle on a sheet of paper). Surface is times. So, the length times the width. That's the surface. So, 3 times 2 is?
Pupil M: 6.
Pupil B: Six of what? Six sheep? Six pigs? (pupil M shakes head) Then what?
Pupil M: 6 meter...6 centimetres?
Pupil S: It has to do with 6.
Pupil B: It is length times width...
Pupil M: Surface...No...6 square meter.
Pupil S: Yes, 6 square meter.

Compared to team A, interactions in team B were generally characterized by less high quality helping behaviour (for instance, giving an explanation—group mean 3.1—applying the received help—group mean 1.8). Their CL motivation scores were also lower; a group average of 3.4.

Pupil B: This is 8.5 and this is 19, ok? So 8 and a half times 2 is 19. So, write 8.5 centimetres here and 19 here, ok? (shows on work sheet of V)

Pupil V: Yes, so I have to write here 8.5 centimetres first.
Pupil B: No no, don't write. You have to do 8.5 centimetres in length and 19 centimetres in width.
Pupil V: Yes... (starts calculating)
Pupil B: But not like that! Look like this (writes on the worksheet of V)
Pupil V: No but...
Pupil B: Look 8.5 centimetres. You have to do it like this, like this (shows by writing on V's worksheet)
Pupil V: Is 8...
Pupil B: No, 8.5. See! (shows on own worksheet)

The excerpts showed that team A used more high quality helping behaviour than team B and their CL motivation was higher. As suggested in the quantitative results section, the linguistic proficiency of team B might have been lower than that of team A. This is backed up by the fact that pupils in team B made less use of full sentences and more use of nonverbal communication (i.e., pointing).

Discussion

In this study, the effect of the teacher's role on the performance of 10–11 year olds in multiethnic elementary classrooms was investigated. It was hypothesized that when the teacher stimulates pupils' use of high quality helping behaviour (experimental group), this results in a larger increase in pupils' math performance than when the teacher lets pupils fend for themselves (control group), especially for pupils with low prior math knowledge and for immigrant pupils. Furthermore, it was expected that pupils in the experimental group experience a higher motivation to cooperate than pupils in the control group, especially pupils with low prior math knowledge and immigrant pupils.

The results partly supported the hypotheses. It was found that national pupils achieved a higher math score in the experimental group than national pupils in the control group. This effect was corroborated at the group level. In addition, post-test math scores of only the teams with low and medium prior math knowledge appeared to be higher in the experimental group. Split for ethnicity the analyses revealed that, in contrast to national teams with high prior math knowledge, the immigrant teams with high prior math knowledge did not score higher in the experimental group than immigrant teams with low prior math knowledge.

With respect to the pupils' motivation to cooperate, it was found that immigrant pupils with high prior math knowledge were more motivated to cooperate when the teacher did not encourage them to use high quality helping behaviour. For immigrant pupils with low prior math knowledge the opposite pattern emerged.

Even though the CL curriculum was of short duration, the teachers did influence the development of pupils' math performance. Unlike earlier studies, for instance by Gillies (2004) and Gillies and Ashman (2000) our study did not find direct support for the hypothesis that the stimulation of pupils' use of high quality helping behaviour by the teachers results in better math performance, but we did find an interaction effect that revealed that immigrant pupils performed better if the teacher did not stimulate their high quality helping behaviour. This finding conflicts with other studies that demonstrate that immigrant pupils' performance is best served when their use of high quality helping behaviour is stimulated (e.g., Webb and Farivar 1994). A study by Kirchmeyer (1993) showed that immigrant pupils were less active when working in ethnically heterogeneous teams. In the present study there was an even distribution of ethnically heterogeneous

teams and teams with only immigrant pupils in the control group. In contrast, in the experimental group the majority of teams were ethnically heterogeneous. Thus, it could be that the presence of national pupils in most teams in the experimental group lowered the activity of the immigrant pupils.

In keeping with the expectations, teams with low prior math knowledge performed better when the teacher stimulated pupils' use of high quality helping behaviour. Moreover, these pupils also were more motivated to cooperate when they were encouraged to use high quality helping behaviour. This latter result is in line with earlier findings (Gillies and Ashman 2000; Johnson and Johnson 2003).

For the teams with high prior math knowledge, the picture was different: the performance of teams with high prior math knowledge was not influenced by whether or not the teacher encouraged pupils to use high quality helping behaviour. Several studies have suggested that pupils who are able to effectively monitor their own learning process need less feedback from the teacher about how they cooperate (e.g., Cohen 1994; Puustinen 1998). Moreover, the motivation of these pupils, typically pupils with high prior math knowledge (Puustinen 1998; Stevens et al. 1991) to cooperate effectively might be undermined when their level of autonomy is restricted (Cohen 1994). We found partial support for this contention: we did find that pupils with high prior math knowledge were more motivated to cooperate in the control group, but this held true for immigrant pupils only. Post-hoc analyses of linguistic proficiency suggested another reason: the higher motivation to cooperate of the immigrant pupils with high prior math knowledge was found to coincide with a lower linguistic proficiency. Further study is warranted to provide more solid support for the claim that the motivation of high ability immigrant pupils to work in highly structured teams is related to their linguistic proficiency.

Some mention must be made of the mixed findings as regards the manipulation check. There was a discrepancy between the teachers' own views and that of the coders. Whereas the coders only detected more discussion of the CL process in favour of the experimental group, the teachers in the experimental group on the other hand indicated that they were more actively teaching high quality helping behaviour during group work. In accordance with other studies this study also suggests that there is a discrepancy between what the teachers think they are capable of with respect to group work and what they are actually doing (Sharan 1990; Vedder and Veendrick 2003).

Limitations

First, classes in our study were randomly assigned to a condition. We did not assign pupils within classrooms randomly to a group. This leaves the possibility that the teacher's educational style might have affected our results. Recent research has shown that the teacher's educational style can be influential in the classroom (e.g., Webb et al. 2006).

Second, the sample was too small to use a statistical multi-level approach. In an attempt to overcome this, analyses were conducted at the individual as well as at the group level. Both levels of analyses yielded a different outcome regarding the role of ethnicity. Whereas at the group level ethnicity added explanatory value to the relationship of the teacher's role and prior math knowledge with math performance, no effect of ethnicity was found at the individual level. This seems a puzzling finding. It has been suggested that individual characteristics such as prior knowledge or the quantity of talk during CL cannot properly account for the learning process at the group level (Barron 2003). In our analyses we interpreted prior knowledge to be individual prior math ability. Barron's study suggests

that team success is best predicted by joint attention to the task at hand and a supportive climate for different ideas. Barron argued that more attention should be paid to interrelational and situated factors, such as the opportunity for positive relational talk, the discussion of ideas, and whether team members feel comfortable with each other. In this study, attention was paid only to prior math knowledge and learning outcomes: no specific attention was paid to process factors like the discussion of ideas. Research has demonstrated this can provide a valuable insight in the mechanisms that drive learning gains (e.g., Kumpulainen and Mutanen 1999; Webb et al. 1995; Wegerif et al. 1999).

Third, mention has to be made of the effect sizes. When we apply Cohen's (1988) criteria, then all effects were small (i.e., equal to or below .20), meaning that there is only a 20% chance that a similar study will yield the same results that we found. A way of enhancing the effect size in a future study might be to implement the CL intervention over a longer period of time. Also, other researchers, like O'Donnell et al. (1990) use "scripts" as a means to structure the peer interactions in groups. The scripts developed by O'Donnell et al. (1990) are protocols that require pupils to alternate between the roles of (either planner/performer or listener/observer). Their study reveals that the use of scripts increased the peer interactions that the pupils display and was related to higher learning gains. Incorporation of scripts in CL programs characterized by teacher's stimulation of pupils' helping behaviour might increase pupils' helping behaviour and their learning gains.

Conclusion

This study demonstrates that, even with a limited amount of time and resources, teachers are able to master at least some of the skills that are needed to successfully carry out group work in multiethnic classes. With more training, teachers may not only become more experienced in the implementation of specific CL skills (like helping behaviour), but also become more aware of their own teaching behaviour during CL. In such a training explicit attention should also be paid to each teacher's unique teaching style.

References

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, *12*, 307–359.

Calderón, M., Hertz-Lazarowitz, R., & Slavin, R. (1998). Effects of bilingual cooperative integrated reading and composition on students making the transition from Spanish to English reading. *The Elementary School Journal*, *99*, 153–165.

Chinn, C. A., O'Donnell, A. M., & Jinks, T. S. (2000). The structure of discourse in collaborative learning. *The Journal of Experimental Education*, *69*, 77–97.

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*, 1–35.

DeVries, D. L., & Slavin, R. E. (1978). Teams-games-tournaments (TTG): Review of ten classroom experiments. *Journal of Research and Development in Education*, *12*, 28–38.

Evers, A., van Vliet-Mulder, J. C., & Groot, C. J. (2000). *COTAN Documentatie van tests en testresearch in Nederland* [Documentation on tests and test research in the Netherlands]. Amsterdam: Boom test uitgevers.

Gijsberts, M. (2004). *Ethnic minorities and integration: Outlook for the future*. The Hague, Netherlands: Social and Cultural Planning Office (SCP).

Gillies, R. M. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction*, *14*, 197–213.

Gillies, R. M., & Ashman, A. F. (1997). Children's cooperative behavior and interactions in trained and untrained work groups in regular classrooms. *Journal of School Psychology, 35*, 261–279

Gillies, R. M., & Ashman, A. F. (2000). The effects of cooperative learning on students with learning difficulties in the lower elementary school. *The Journal of Special Education*, *34*, 19–27.

Hijzen, D., Boekaerts, M., & Vedder, P. (2006). The relationship between the quality of cooperative learning, students' goal preferences and perceptions of contextual factors in the classroom. *Scandinavian Journal of Psychology*, *47*, 9–21.

Janssen, J., Kraemer, J. M., & Noteboom, A. (1996). *Rekenen-Wiskunde 3. Handleiding* [Arithmetic-Math 3. Manual]. Arnhem: Citogroep.

Johnson, D. W., & Johnson, F. P. (2003). *Joining together, group theory and group skills* (8th ed.). Boston: Pearson Education, Inc.

Kirchmeyer, C. (1993). Multicultural task groups: An account of the low contribution level of minorities. *Small Group Research*, *24*, 127–148.

Klingner, J. K., Vaughn, S., & Schumm, J. S. (1998). Collaborative strategic reading during social studies in heterogeneous fourth-grade classrooms. *The Elementary School Journal*, *99*, 3–22.

Kumpulainen, K., & Mutanan, M. (1999). The situated dynamics of peer group interaction: An introduction to an analytic framework. *Learning and Instruction*, *9*, 449–473.

Mulryan, C. M. (1994). Perceptions of intermediate students' cooperative small-group work in mathematics. *Journal of Educational Research, 87*, 280–291.

O'Donnell, A. M., Dansereau, D. F., Hall, R. H., Skaggs, L. P., Hythecker, V. I., Peel, J. L., & Rewey, K. L. (1990). Learning concrete procedures: Effects of processing strategies and cooperative learning. *Journal of Educational Psychology, 82*, 171–177.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*, 61–86.

Puustinen, M. (1998). Help-seeking behavior in a problem-solving situation: Development of self-regulation. *European Journal of Psychology of Education*, *13*, 271–282.

Qin, Z., Johnson, D. W., & Johnson, R. T. (1995). Cooperative versus competitive efforts and problem solving. *Review of Educational Research*, *65*, 129–143.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, *95*, 240–257.

Saleh, M., Lazonder, A. W., & de Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science*, *33*, 105–119.

Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. In P.D. Pearson, & A. Iran-Nejad (Eds.), *Review of research in education* (pp. 1–24). Washington DC: American Educational Research Association.

Sharan, S. (1990). Cooperative learning and helping behavior the multi-ethnic classroom. In H. C. Foot, M. J. Morgan, & R. H. Shute (Eds.), *Children helping children* (pp. 151–176). Chichester: John Wiley.

Slavin, R. E., & Cooper, R. (1999). Improving intergroup relations: Lessons learned from cooperative learning programs. *Journal of Social Issues*, *55*, 647–663.

Stevens, R. J., Slavin, R. E., & Farnish, A. M. (1991). The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology*, *83*, 8–16.

Tesser, P. T. M., & Iedema, J. (2001). *Rapportage Minderheden 2001* [Minorities Report 2001]. SCP: Den Haag.

Vedder, P., & Veendrick, A. (2003). The role of the task and reward structure in cooperative learning. *Scandinavian Journal of Educational Research*, *47*, 529–542.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press (Original work published in 1930).

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal*, *31*, 369–395.

Webb, N. M., Nemer, K. M, & Ing, M. (2006). Small group reflections: Parallels between teacher discourse and student behavior in peer-directed groups. *The Journal of the Learning Sciences*, *15*, 63–119.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Cafree (Eds.), *Handbook of educational psychology* (pp. 841–873). New York: Simon & Shuster MacMillan.

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology*, *87*, 406–423.

Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: an empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction*, *9*, 493–516.