

The effect of stimulating immigrant and national pupils' helping behaviour during cooperative learning in classrooms on their maths-related talk

Michiel Bastiaan Oortwijn*, Monique Boekaerts and Paul Vedder

Center for the Study of Education and Instruction, Leiden University, The Netherlands

This study examined whether stimulation of immigrant and national pupils' use of high-quality helping behaviour (experimental condition) during cooperative learning (CL) in classrooms boosts their maths-related talk more than in an educational situation in which such stimulation is largely absent (control condition). A total of 59 elementary-age pupils enrolled in a CL maths curriculum of 11 lessons. They were video taped during two lessons while working together on maths assignments to assess their maths-related talk. We found that the quality of maths-related talk was higher in the experimental condition. Furthermore, immigrant pupils' used less maths-related talk than the national pupils. Implications are discussed.

Keywords: cooperative learning; helping behaviour; maths-related talk; immigrant pupils; elementary schools

Introduction

Helping behaviour plays an important role in the classroom. A study by Newman and Schwager (1993) has demonstrated that pupils often are inclined to ask the teacher for help, since they view the teachers as more capable to facilitate learning than their peers. This is unfortunate since research demonstrates that stimulating peers to help one another can augment their learning gains more when compared to pupils who are not stimulated to help one another (e.g. Gillies and Ashman 2000). However, this is dependent on the quality of the help. For instance, Webb, Nemer and Ing (2006) have found that stimulating pupils to use low-quality help negatively affects their learning gains. Noreen Webb and her colleagues (e.g. Webb and Farivar 1994; Webb, Farivar and Mastergeorge 2002) have demonstrated that only high-quality helping behaviour is positively related to learning gains. Webb defined high-quality help as help that is sufficiently elaborated, correct, on time, links up to the need for help and lets the help receiver apply the help that is given. Webb and Farivar (1994) have shown that the teacher can positively affect pupils' use of high-quality helping behaviour. High-quality helping behaviour is defined here as helping behaviour that is characterised by asking, giving and applying explanations. In this study, we investigated how the stimulation of high-quality helping behaviour affected pupils' linguistic proficiency.

Studying linguistic proficiency by investigating peer interactions

Increasingly, researchers are recognising that peer interactions are an essential force that drives students' cognitive development in CL settings (e.g. Gillies 2004; Gillies and

*Corresponding author. Email: oortwijn@fsw.leidenuniv.nl

Ashman 2000; Keefer, Zeitz and Resnick 2000; Webb, Farivar and Mastergeorge 2002; Wegerif, Mercer and Dawes 1999). Most of the aforementioned studies have focused on different verbal aspects of peer interactions to assess its relationship with cognitive development. For instance, Webb, Farivar and Mastergeorge (2002) have found a positive relationship between helping behaviour and mathematics performance. Wegerif, Mercer and Dawes (1999) have provided evidence for a positive association between the occurrence of group discussions during peer interactions and subsequent cognitive development. These researchers all have provided empirical support for the notion that peer interactions are related to cognitive growth. But there are differences in opinion about which part of the peer interaction process positively influences cognitive development.

Measuring linguistic development

Researchers like Zack and Graves (2001) and Hicks (1995) have explored the relation between language and cognitive development in a CL setting. Although these explorative studies have provided important insights, they have solely focused on the process of interacting. In contrast, our main interest in this study was to investigate whether we could enhance pupils' linguistic proficiency, in particular that of immigrant pupils, by stimulating the pupils to use high-quality helping behaviour during CL.

We investigated pupils' linguistic development using two approaches. The first approach regards the frequency with which pupils use specific signalling words, like question markers or words that are associated with meta-linguistic and mental activities, like "but" and "because" (cf. Vedder, Kook and Muysken 1996). The second approach regards the investigation of a set of semantically linked words within a specific context. That is, the focus lies on the comprehension of figurative speech, also known as idioms (e.g. Cain, Oakhill and Lemmon 2005). Research into pupils' use of idioms assesses the degree to which students take idioms literally. For instance, a study might focus on whether pupils understand that when someone "hits the road", this does not mean that this person is going to punch the road, but rather that he or she is leaving. Research has demonstrated that especially pupils with linguistic problems (i.e. immigrant pupils in The Netherlands) are having difficulties with the correct interpretation of figurative speech (Levorato and Cacciari 1995).

We combined elements from both approaches in the context of a CL mathematics curriculum. We hypothesised that understanding the mathematical meaning of particular maths concepts (like "circumference") facilitates maths gains. Following Niemi (1996), we defined understanding in two ways. Firstly, we distinguished *semantic* understanding. We interpreted this as the pupils' explicit knowledge of the specified maths concepts. That is, correctly verbalising the definition. Secondly, we distinguished *general* understanding, which is a more implicit grasp of the conceptual framework in which the specified maths concepts are embedded. More specifically, general understanding was interpreted here as the "operationalisation" of the concept in a specific context.

Performance of immigrant pupils in Dutch elementary schools

In The Netherlands there are three major discernible ethnic groups with respect to their performance at school: (1) Moroccan, Turkish and Antillean youth, (2) Surinamese and other ethnic youth groups (e.g. Asian, former Yugoslavia) and (3) Dutch youth. Tesser and Iedema (2001) have shown that especially the performance of the Moroccan, Turkish and Antillean groups falls behind; for instance, their linguistic performance is delayed by two years in the last grade of the elementary school when compared to that of national

pupils. Since these ethnic groups together form the majority of immigrant youth, it may come as no surprise that immigrant pupils perform below the national mean, both linguistically and mathematically (Bosker and Guldemond 2004; Tesser and Iedema 2001). Not only do teachers have insufficient resources to cope with the highly divergent language abilities of the immigrant pupils, but due to the high communicative requirements of most modern maths curricula, they also lack the skills to prevent immigrant pupils' maths performance from falling behind as compared to the national mean. This is translated into a significant deprivation in maths and linguistic performance of the immigrant pupils at the end of elementary school (Tesser and Iedema 2001).

A US study by Webb and Farivar (1994) has demonstrated that the positive effect of the stimulation of high-quality helping behaviour during CL on performance was particularly salient for the immigrant pupils (i.e. educationally disadvantaged Latino and African American pupils) (see also Calderón, Hertz-Lazarowitz and Slavin 1998).

Research design and hypotheses

We created two conditions to investigate the effect of high-quality helping behaviour on the maths-related talk of both national and immigrant pupils in elementary schools: an experimental condition (in which the pupils were stimulated to use high-quality helping behaviour) and a control condition (in which the pupils were not stimulated to use high-quality helping behaviour). Our hypotheses were:

- (1) Being in the experimental condition stimulates pupils' maths-related talk, that is the frequency of question markers, conjunctions and the quality of mathematical idioms, more than being in the control condition.
- (2) Being in the experimental condition reduces the difference in maths-related talk between immigrant pupils and national pupils more than being in the control condition.

Method

Sample characteristics

Video recordings were made of both teacher–pupil interactions (during two randomly selected lessons to check the treatment integrity) and the peer interactions. Regarding the peer interactions, video recordings were made of 15 groups (59 pupils; mean age 134.3 months, SD 6.3 months), randomly selected from eight fifth-grade elementary classes (see Table 1). In the control condition there were nine groups, comprising 35 pupils (mean age = 133.4 months, SD = 5.9; 16 male, 19 female, 8 Dutch, 27 immigrant). There were six groups in the experimental condition, comprising 24 pupils (mean age = 135.5 months, SD = 6.9; 12 male, 12 female, 11 Dutch, 13 immigrant). With respect to ethnicity, pupils were defined as national if both parents were Dutch and defined as immigrant when one or both parents had a non-Dutch nationality. In the control condition, 85% of the immigrant pupils were Turkish or Moroccan. In the experimental condition, 69% of the immigrant pupils were Turkish or Moroccan. The average length of the video recordings was 941.1 seconds (SD = 229.0) and did not differ between the experimental and the control condition.

Procedure

The CL curriculum consisted of 11 CL mathematics lessons. In the first two CL lessons, the teacher instructed all the pupils how to use particular CL rules. These were “Everyone

Table 1. Sample characteristics.

Group	No. of groups	No. of pupils	Mean age (months) (SD)	Ethnicity	Gender
1	9	35	133.4 (5.9)	8 Dutch	16 male
				27 immigrant	19 female
2	6	24	135.5 (6.9)	11 Dutch	12 male
				13 immigrant	12 female

cooperates”, “Everyone listens to each other”, “Everyone shares their knowledge and opinions”, and “Checks whether everyone agrees”, “Ask precise questions”, “Continue asking in case of ambiguities”, “Think before asking a question”, “Ask for help on time”, “Fine-tuning of the level of guidance to the need for help that is requested”, “Giving a clear and precise answer”, “Giving the help-receiver a chance to apply the help given”, “Continuing to ask if the question for help is unclear” and “Giving help when needed”. The CL rules were practised and written on a poster, which was then displayed at a spot that was clearly visible to the children.

After the two CL training lessons (similar for both the experimental and the control condition), all pupils received the maths curriculum of nine, one hour CL lessons. Each lesson consisted of two open-ended authentic maths assignments with a narrative structure. The maths assignments used in this study dealt with area, scale, fractions, percentage and circle diagrams. All assignments were adjusted for CL purposes using authentic maths assignments from a regular maths curriculum. Authentic maths assignments are mathematical tasks with a strong narrative structure and which are embedded in contexts familiar to the children, such as calculating the area of the classrooms in their school. After the CL maths curriculum, the pupils individually completed a maths exam.

Manipulation

The teachers were randomly assigned to either the control or the experimental condition. Teachers in the control condition were instructed to guide the group work only if pupils: (1) talked too loudly (disturbed other groups), (2) did not listen to one another and (3) made fun of one another.

We instructed the teachers in the experimental condition to promote the pupils’ use of high-quality helping behaviour. Additionally, the teachers in the experimental condition discussed the pupils’ use of high-quality helping behaviour with the whole class at the end of each lesson. The teachers were instructed to stimulate both the use of high-quality helping behaviour of the individual pupils and the groups. Detailed protocols aided the teachers to implement CL (experimental and control condition).

In six 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. There were no differences between conditions in the number of years of teachers’ experience.

Instruments

Coding maths-related talk

Overview of the coding scheme. The coding scheme consisted of three dimensions. Dimension 1 assessed the frequency of: (a) question markers, “what” question markers

(value 1) and “why” question markers (value 2); and (b) conjunctions, like “because” and “unless”. Dimension 2 concerned the frequency and quality of understanding of the mathematical concepts “scale”, “area”, “circumference” and “estimation”. For each concept, two levels were distinguished: (1) low-level application of the mathematical concept (inappropriate use of a definition, verbalisation of a maths concept only, use of numbers only); and (2) high-level application of a mathematical concept (use of a context definition, sharing a definition by two pupils or use of an abstract definition). Dimension 3 assessed the frequency of the occurrence of a number of mathematical words, which could aid in the problem-solving process (e.g. “fraction”, “divide”).

Coding procedure. Two coders scored the video recorded peer interactions on the computer, using the software program Observer 5.0 (Noldus 2003). With this program, the relevant utterances that were made by the pupils were marked by the coders on a timeline as they appeared in the video file in “real life” (i.e. there was no need to transcribe the interactions). While coding, the coders had at their disposal comprehensive coding instructions, the maths assignments that the pupils worked on and a list with all the correct problem-solving steps and the right solutions for the assignments. Both coders (i.e. the first author and a student who was unaware of the experimental manipulation) practised the coding scheme by coding several video taped peer interactions (which were not used in the analyses) and discussing differences in scores until agreement was reached. After this training, the inter-coder reliability was calculated on approximately 20% of the data (six video recordings). For dimension 1(a) kappa was 0.79, for 1(b) the inter-coder agreement was 84%, kappa was 0.62. For dimension 2, inter-coder agreement was 93%, kappa was 0.62. For dimension 3, inter-coder agreement was 94%, kappa was 0.93. The second coder then coded the remaining recordings.

Prior linguistic proficiency and prior maths performance

To assess prior linguistic proficiency and mathematic performance, we used the linguistic and mathematical subtests from the national testing service, CITO (Janssen, Kraemer and Noteboom 1996). All pupils completed the linguistic and mathematic subtests prior to the CL curriculum. We used the scores to assess the baseline linguistic and mathematic performance of all pupils.

Linguistic proficiency. We used the sum scores of the two dimensions “vocabulary” and “reading comprehension” from this subtest. Research has revealed that this subtest has a good internal consistency, $\alpha = 0.80$ (Evers, Van Vliet-Mulder and Groot 2000). We averaged the sum scores of both these two dimensions into a new variable, which we labelled “prior linguistic proficiency”.

Maths pre-test. Scores from a curriculum-independent maths test (CITO; Janssen, Kraemer and Noteboom 1996) were used to assess the baseline maths performance of all pupils. A Pearson’s correlation test showed that the pre-test significantly correlated with the post-test, $r = 0.86, p < 0.001$. 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 maths test has a good reliability, $\alpha = 0.94$ (Evers, Van Vliet-Mulder and Groot 2000).

Treatment integrity

We checked the treatment integrity with two instruments. (1) *Teacher checklist on stimulation of high-quality helping behaviour*. The teachers completed a checklist at the end of every other lesson, on which they indicated on a four-point Likert scale (1 = “very often” and 4 = “very little”) (a) to what extent they had implemented the general CL rules, and (b) their stimulation of pupils’ high-quality helping behaviour during the last CL lesson. The items of the checklist corresponded to the CL instructions in the lesson-to-lesson protocol for the experimental condition. (2) *Video taped teacher–pupil interactions*. All teachers were video taped during two or more lessons as they provided feedback to the pupils to check whether the frequency and quality of the teacher–pupil interactions that were related to the stimulation of pupils’ high-quality helping behaviour differed between the conditions.

The coding scheme consisted of two dimensions, comprising 14 items. The first dimension (six items, $\alpha = 0.71$) was about the presence of whole-class reflection on the group work (e.g. “Does the teacher reflect on the group performance?”). The second dimension (eight items, $\alpha = 0.86$) consisted of items concerning the teacher’s whereabouts and activities during group work (e.g. “Does the teacher encourage the group members to ask each other questions?”). Cohen’s kappa for two recordings (approximately 10%) was 0.73 for dimension 1 and 0.62 for dimension 2.

Results

We started the analyses by assessing pupils’ prior linguistic proficiency, then checked the manipulation integrity. We proceeded with the analyses of the relation between the stimulation of the pupils’ use of high-quality helping behaviour and their maths-related talk and whether ethnicity interacted with this relation.

Prior linguistic proficiency

The immigrant pupils had a lower prior linguistic proficiency (mean 2.36, SD = 0.72) than national pupils (mean 2.97, SD = 0.82), $t(57) = -2.76$, $p < 0.01$, Cohen’s $d = 0.79$. We found no differences between the two conditions in prior linguistic proficiency.

Manipulation integrity

Regarding the teacher checklist on the stimulation of high-quality helping behaviour, we found that teachers in the experimental condition reported instructing pupils significantly more in the use of helping skills, $t(21) = -3.37$, $p < 0.005$, Cohen’s $d = 1.48$, than the teachers in the control condition. We did not find differences between the two conditions on the dimensions “General social rules”.

Analysis of the coded video tapes of teacher–pupil interactions revealed that teachers in the experimental condition reflected more on the group work than teachers in the control condition, $t(16) = -1.78$, $p < 0.05$, Cohen’s $d = 0.58$. We did not find differences between the two conditions for the dimension “Feedback on group work during CL”.

Relationship of high-quality helping behaviour with maths-related talk

We applied a logarithmic transformation to all variables, since the data were substantially skewed, and there was a significant heterogeneity of variance between the conditions (see

also Tabachnick and Fidell 2001). After this, the kurtosis and skewness divided by their standard errors were smaller than 2 for all variables, which is acceptable (Tabachnick and Fidell 2001). We carried out a series of 2(condition) \times 2(ethnicity) univariate analyses of covariance to assess their relationship with the categories of the coding scheme that assessed maths-related talk (“High-quality question markers”, “Low-quality question markers”, “Conjunctions”, “Low-level understanding of maths concepts” – only seven of the 2041 coded utterances regarded “High-level understanding of maths concepts” and, therefore, they were left out of the analysis – and “Use of mathematical words”). “Prior linguistic proficiency” and “Maths pre-test” formed the covariate. Using a Bonferroni correction, the p -value was set at 0.01. The analyses yielded the following results.

There were significant main effects for “Condition” with “High-quality question markers”, $F(1,42) = 9.55, p < 0.005 [\eta^2 = 0.19]$, explaining 19% of the variance, “Low-quality question markers”, $F(1,52) = 9.10, p < 0.005 [\eta^2 = 0.15]$, explaining 16% of the variance, and “Use of mathematical words” $F(1,40) = 14.89, p < 0.001 [\eta^2 = 0.27]$, explaining 27% of the variance. Pupils who were stimulated to use high-quality helping behaviour used more mathematical words, and both low- and high-quality question markers.

Additionally, we found a significant main effect for “Ethnicity” with “High-quality question markers”, $F(1,42) = 8.40, p < 0.007 [\eta^2 = 0.16]$, explaining 17% of the variance. Immigrant pupils used less high-quality question markers than national pupils.

Analyses at the group level

Since we focused on the pupils’ maths-related talk, the individual pupils were the unit of analysis. Nevertheless, as group work was the focus of research in this study, an explorative analysis of the mathematics 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 multi-level approach. Inspired by Gillies and Ashman (2000), Webb and Farivar (1994) and Saleh, Lazonder and de Jong (2005), analyses at the group level were executed by aggregating pupils’ scores on the dimensions of maths-related talk. Because of the small n , we carried out a Mann-Whitney test to cross-validate the relation of “Condition” with “High-quality question markers”, “low-quality question markers” and “Use of mathematical words” we found at the individual level. We found that there was a trend for pupils to use more high-quality question markers in the experimental condition (10.25) than in the control condition (6.50), $Z(15) = -1.60, p = 0.056$, to use more low-quality question markers in the experimental condition (11.67) than in the control condition (5.56), $Z(15) = -2.60, p < 0.005$ and more mathematical words in the experimental condition (10.67), when compared to the control condition (6.22), $Z(15) = -1.90, p < 0.03$.

Interaction (effect) of condition and ethnicity on (scores for) maths-related talk

Using a p -value of 0.01, we did not find an interaction between ethnicity and condition.

Discussion

In this study, we have investigated whether stimulating immigrant and national pupils’ use of high-quality helping behaviour augments their use of specific question markers, conjunctions and of mathematical concepts and words. Also, we have examined if being in the experimental condition reduces the difference in maths-related talk between immigrant pupils and national pupils more than being in the control condition.

We found that stimulating pupils to give and receive help positively affected their use of maths-related words and questions, both high-quality and low-quality. This result extends earlier findings (Webb and Farivar 1994; Webb, Farivar and Mastergeorge 2002). Stimulating pupils' use of high-quality helping behaviour did not raise the frequency of mathematical idioms. Possibly this was due to the overall low use of mathematical idioms by the pupils. Two reasons are put forward here for the possible low use of mathematical idioms by the pupils. The first is methodological: the inclusion criteria for mathematical idiom were too rigid. Although pupils frequently verbalised a numerical rule without reference to a mathematical concept, these verbalisations usually were too ambiguous to make accurate coding possible. For instance, pupils frequently multiplied two numbers while working on area and scale problems. Nevertheless, it was sometimes unclear whether they were (incorrectly) calculating the scale, or whether they were calculating an area.

The second reason put forward here is that pupils are just not accustomed to explicitly labelling the mathematical operation when referring to a specific mathematics concept in their peer interactions (see also Bennett and Dunne 1991). This cannot be caused by inexperience with the use of such concepts, since the pupils did make frequent use of other maths-related talk, like mathematical words. Another explanation is suggested by Cain, Dakhill and Lemmon (2005), who argued that there are two approaches to understanding idioms: through semantic analysis or by making use of the context. Perhaps the pupils in our study, although they did make use of the context in order to work with the mathematical idioms, had difficulty verbalising implicit knowledge. This was noticeable on a number of recordings. In some groups, different pupils simultaneously used definitions of both area and circumference when interacting about how to calculate the area of a classroom, without correcting one another. This suggests that, in spite of a shared understanding of the mathematical concept, evidenced by correct mathematical solutions, pupils still seemed to have problems to relate this implicit knowledge to the right mathematical idiom.

As mentioned before, pupils in the experimental condition used both more low- and high-quality question markers than pupils in the control condition. How might this apparent contradiction be explained? Two possible explanations are discussed here. The first explanation is that the higher frequency of high-quality question markers neutralised the adverse effect of the low-quality question markers. Although not directly related to post-test maths performance, the high-quality question markers could have affected the relationship of low-quality question markers with post-test maths performance. There is a large body of literature which suggests that high-quality questioning is positively related to learning gains (Fantuzzo et al. 1989; King 2002). A larger sample is required to replicate the positive effect of high-quality question markers with subsequent performance.

A second possible explanation is that a functional differentiation occurred: it might be that in the experimental condition, the low-quality question markers were used more for the management of the groups' CL process and the high-quality question markers were used more for the maths-related problem-solving process. There is no direct evidence for this, but the results did show that pupils used more mathematical words in the experimental condition compared to the control condition. This suggests that the pupils in the experimental condition talked more about maths.

Related to this, we found that immigrant pupils used less high-quality question markers than national pupils. This is in line with other studies that also have revealed that immigrant pupils contribute less to interactions in groups (Kirchmeyer 1993; Oortwijn et al. 2008).

We could not corroborate our prediction that the difference between immigrant and national pupils in their use of maths-related talk in the experimental condition was smaller when compared to the control condition. It suggests that ethnicity alone does not predict the

use of maths-related talk: other aspects may be stronger predictors of maths-related talk, such as the parents' socio-economic status (Gijsberts 2004) and motivation to work in groups (Oortwijn, Boekaerts and Vedder 2008).

Limitations

Three limitations are mentioned here. First of all, the sample was relatively small. A larger sample size is essential to corroborate the findings reported in this paper. Second, there were more immigrant pupils in the control condition than in the experimental condition. This may have distorted some of our findings. Third, the effect sizes were generally of modest size (Cohen 1977). This means that we should be careful as to the possible significance of the findings for the daily classroom practice: on the basis of our findings, we cannot firmly argue that CL has an educational value in the reduction of immigrant pupils' linguistic problems, at least in the short term. There are reasons to suspect that a longer intervention increases the effect size, and thus the educational value of CL (Webb, Nemer and Ing 2006).

Conclusion

Earlier studies have shown that peer helping behaviour is positively associated with pupils' learning gains (Gillies and Ashman 2000; Webb and Farivar 1994). What these studies also illuminate is that in order for pupils to help one another effectively, they have to be guided by the teacher. In line with this, our study suggests that learning in the CL context requires that the pupils' peer interactions are structured to maximise the development of their maths related talk: pupils whose peer interactions are not structured resort to low-quality interaction patterns.

Notes on contributors

Michiel Oortwijn is an assistant professor in the department of education and child studies of Leiden University. His current research focuses on the relation of interethnic relations and attitudes with problem behaviour and drop-out.

Monique Boekaerts holds the chair of Learning and Instruction. She has played a leading role in the development of the Motivation construct, designing the first situation-specific measurement instrument to assess motivation in the classroom (The On-line Motivation Questionnaire). Her theories and empirical studies are also of high importance for the international recognition of the Self-regulation construct.

Paul Vedder holds chairs in clinical child and youth care and in cultural diversity and learning in the department of education and child studies of Leiden University. His research focuses on social competence, multilingualism and the quality of interethnic relationships.

References

- Bennett, N., and E. Dunne. 1991. The nature and quality of talk in co-operative classroom groups. *Learning and Instruction* 1: 103–18.
- Bosker, R.J., and H. Guldemond. 2004. *Een herijking van de gewichtenregeling [A recalibration of the financial arrangements for students with an educational deprivation]*. Groningen: GION.
- Cain, K., J. Oakhill, and K. Lemmon. 2005. The relation between children's reading comprehension level and their comprehension of idioms. *Journal of Experimental Child Psychology* 90: 65–87.
- Calderón, M., R. Hertz-Lazarowitz, and R. Slavin. 1998. Effects of bilingual cooperative integrated reading and composition on students making the transition from Spanish to English reading. *Elementary School Journal* 99: 153–65.
- Cohen, J. 1977. *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

- Evers, A., J.C. Van Vliet-Mulder, and C.J. Groot. 2000. *COTAN documentatie van tests en testresearch in Nederland* [COTAN documentation on tests and test research in The Netherlands]. Amsterdam: Boom test uitgevers.
- Fantuzzo, J.W., R.E. Riggio, S. Connelly, and L.A. Dimeff. 1989. Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. *Journal of Educational Psychology* 81: 173–7.
- Gijsberts, M. 2004. *Ethnic minorities and integration: Outlook for the future*. The Hague: Social and Cultural Planning Office.
- 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., and A.F. Ashman. 2000. The effects of cooperative learning on pupils with learning difficulties in the lower elementary school. *Journal of Special Education* 34: 19–27.
- Hicks, D. 1995. Discourse, learning and teaching. In *Review of Research in Education 21*, ed. M.W. Apple, pp. 49–95. Washington, DC: American Educational Research Association.
- Janssen, J., J.M. Kraemer, and A. Noteboom. 1996. *Rekenen-wiskunde 3: Handleiding* [Arithmetic – maths 3: Manual]. Arnhem: Citogroep.
- Keefer, M.W., C. M. Zeitz, and L.B. Resnick. 2000. Judging the quality of peer-led student dialogues. *Cognition and Instruction* 18: 53–81.
- King, A. 2002. Structuring peer interaction to promote high-level cognitive processing. *Theory into Practice* 41: 33–9.
- Kirchmeyer, C. 1993. Multicultural task groups: An account of the low contribution level of minorities. *Small Group Research* 24: 127–48.
- Lavorato, M.C., and C. Cacciari. 1995. The effects of different tasks on the comprehension and production of idioms in children. *Journal of Experimental Child Psychology* 60: 261–83.
- Newman, R.S., and M.T. Schwager. 1993. Students' perceptions of the teacher and classmates in relation to reported help seeking in math class. *Elementary School Journal* 94: 3–17.
- Niemi, D. 1996. Assessing conceptual understanding in mathematics: Representations, problem solutions, justifications and explanations. *Journal of Educational Research* 89: 351–63.
- Noldus L.P.J.J. 2003. *The observer (Version 5.0)* [Computer software]. Wageningen: Noldus Information Technology Inc.
- Oortwijn, M.B., M. Boekaerts, and P. Vedder. 2008. The impact of the teacher's role and pupils' ethnicity and prior knowledge on pupils' performance and motivation to cooperate. *Instructional Science* 36, no. 5: 251–68.
- Oortwijn, M.B., M. Boekaerts, P. Vedder, and J.W. Strijbos. 2008. Helping behaviour during cooperative learning and learning gains: The role of the teacher and of pupils' prior knowledge and ethnic background. *Learning and Instruction* 18, no. 2: 146–59.
- Saleh, M. A.W. Lazonder, and T. de Jong. 2005. Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science* 33: 105–19.
- Tabachnick, B.G., and L.S. Fidell. 2001. *Using multivariate statistics*. 4th edn. Boston MA: Allyn & Bacon.
- Tesser, P.T.M., and J. Iedema. 2001. *Rapportage minderheden 2001* [Minorities report 2001]. The Hague: SCP.
- Vedder, P., H. Kook, and P. Muysken. 1996. Language choice and functional differentiation of languages in bilingual parent–child reading. *Applied Psycholinguistics* 17: 461–84.
- Webb, N.M., and S. Farivar. 1994. Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal* 31: 369–95.
- Webb, N.M., S.H. Farivar, and A.M. Mastergeorge. 2002. Productive helping in cooperative groups. *Theory into Practice* 41: 13–20.
- Webb, N.M., K.M. Nemer, and M. Ing. 2006. Small group reflections: Parallels between teacher discourse and student behavior in peer-directed groups. *Journal of the Learning Sciences* 15: 63–119.
- Wegerif, R., N. Mercer, and L. Dawes. 1999. Children's talk and the development of reasoning in the classroom. *British Educational Research Journal* 25: 95–111.
- Zack, V., and B. Graves. 2001. Making mathematical meaning through dialogue: “Once you think of it, the Z minus three seems pretty weird”. *Educational Studies in Mathematics* 46: 229–71.

Copyright of Educational Studies (03055698) is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Copyright of Educational Studies (03055698) is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.