The effect of productive classroom talk and metacommunication on young children's oral communicative competence and subject matter knowledge: An intervention study in early childhood education

van der Veen, C.; de Mey, J.R.P.B.; van Kruistum, C.J.; van Oers, B.

published in
Learning and Instruction
2017

DOI (link to publisher)
10.1016/j.learninstruc.2016.06.001

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
van der Veen, C., de Mey, J. R. P. B., van Kruistum, C. J., & van Oers, B. (2017). The effect of productive classroom talk and metacommunication on young children's oral communicative competence and subject matter knowledge: An intervention study in early childhood education. *Learning and Instruction*, *48*, 14-22. https://doi.org/10.1016/j.learninstruc.2016.06.001

Learning and Instruction 48 (2017) 14–22

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

The effect of productive classroom talk and metacommunication on young children's oral communicative competence and subject matter knowledge: An intervention study in early childhood education

Chiel van der Veen*, Langha de Mey, Claudia van Kruistum, Bert van Oers

Department of Research and Theory in Education, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 23 October 2015
Received in revised form
29 April 2016
Accepted 1 June 2016
Available online 26 June 2016

Keywords:
Productive classroom dialogue
Oral communicative competence
Metacommunication
Subject matter knowledge
Early childhood education
Multilevel analyses

ABSTRACT

The aim of the present study was to investigate the effect of productive classroom talk and metacommunication on the development of young children's oral communicative competence and subject matter knowledge. This study can be characterized as a quasi-experimental study with a pre-test-intervention-post-test design. A total of 21 teachers and 469 children participated in this study. 12 teachers were assigned to the intervention condition and participated in a Professional Development Program on productive classroom dialogue. Multilevel analyses of children's oral communicative competence pre- and post-test scores indicated that our intervention had a significant and moderate to large effect on the development of young children's oral communicative competence. No significant effects were found for children's subject matter knowledge. The results of this study suggest that dialogically organized classroom talk is more beneficial than non-dialogical classroom talk for the development of children's oral language skills.

1. Introduction

In the past decades, research on classroom discourse has received a considerable amount of attention in the educational sciences and beyond (for example, Cazden, 2001; Howe & Abedin, 2013; Resnick, Asterhan, & Clarke, 2015). This body of research has provided some support for the claim that dialogically orchestrated classroom talk is positively related to children's (content) learning and development (Kiemer, Gröschner, Pehmer, & Seidel, 2015; Nystrand & Gamoran, 1991; Nystrand, Wu, Gamoron, Zeiser, & Long, 2003; Wegerif, Mercer, & Dawes, 1999). However, according to a recent review on classroom discourse by Howe and Abedin (2013) there is a lack of empirical evidence to decide "whether certain modes of organization are more beneficial than others" (p. 325) as most of the studies on classroom discourse are qualitative in nature. Howe and Abedin suggest, "it is time to take risks" (p. 346) and complement this body of research with quantitative studies. Furthermore, most studies on classroom discourse are conducted in the upper grades of primary education or in secondary education schools. Much remains unknown about the potentials of dialogic classroom talk in early childhood classrooms. Besides, most studies on classroom discourse focus on dialogue in small-group contexts. As many teachers struggle to orchestrate classroom discussion in whole-class settings, it might be worthwhile to investigate how dialogic practices can be transferred to whole-class contexts. Finally, many studies that report on the relation between classroom discourse and children's learning or development only take measurements of children's subject matter knowledge or their reasoning skills into account. Much remains unknown about the benefits of dialogically organized classroom talk for the development of children's communication or pragmatic language skills. In this article, therefore, we will focus on the possibilities of dialogically organized classroom talk in whole-class settings – which we will refer to as *productive classroom dialogue* (van der Veen, van Kruistum, & Michaels, 2015) – for the development of young children's oral communicative competence.

* Corresponding author.
E-mail addresses: chiel.vander.veen@vu.nl (C. van der Veen), j.r.p.b.de.mey@vu.nl (L. de Mey), c.j.van.kruistum@vu.nl (C. van Kruistum), bert.van.oers@vu.nl (B. van Oers).

http://dx.doi.org/10.1016/j.learninstruc.2016.06.001
0959-4752/

1.1. Productive classroom dialogue

The concept of productive classroom dialogue (PCD) is rooted in Sarah Michael's and Cathy O'Connor's (2012; 2015) work on classroom talk. In PCD "teachers aim to break away from interaction patterns that are predominantly teacher-steered and based on recitation" (van der Veen et al., 2015, p. 321) and to move beyond the well-known and often used Initiation-Response-Evaluation (IRE) sequence (see Cazden, 2001; see also the examples in Table 3). IRE sequences follow a recitation format in which the teacher poses a closed question, one student (or multiple students in chorus) gives a short response, and the teacher evaluates the correctness of this response. In PCD, however, children are given space to say more, carefully listen to one another, reason, and think together (Michaels & O'Connor, 2012). It aims for a classroom culture in which children take each other seriously and move beyond the boundaries of their own thinking and understanding. Furthermore, it intends to improve children's understandings and communicative abilities by collaboratively producing and negotiating new ideas.

1.1.1. Productive talk moves

In PCD, teachers have several talk moves at their disposal that can be used to orchestrate classroom talk. These talk moves can be seen as tools (Michaels & O'Connor, 2015) that teachers can use to encourage children to (a) share, expand or clarify their initial ideas (*Can you say more about it?*) (b) listen to one another and take other's ideas seriously (*Who think they understood what X said and can put it into their own words?*), (c) deepen their reasoning (*Why do you think that?*), and (d) think with each other and build on each other's ideas (*Can you add onto his idea? Do you agree/disagree? Why?*). The use of these talk moves has turned out to be strongly related to children's academic learning (O'Connor, Michaels, & Chapin, 2015) as they open up the conversation and support teachers to move beyond Initiation-Reponses-Evaluation sequences. A detailed overview of the talk moves used in this study can be found in Appendix A.

1.1.2. Metacommunicative moves

In our work on classroom talk, we have added a category to Michaels' and O'Connor's set of talk moves: the metacommunicative moves. These moves are pedagogical strategies that teachers can use to explicitly reflect on children's communicative performance; they support talk about talk (see Appendix A). The importance of the use of metacommunication in young children's classroom talk derives from the work of Robinson and Robinson (1983). In their experiments, they showed that young children are often unaware of ambiguities and misunderstandings in communication. Robinson and Robinson (1982; 1983) found that making non-comprehension explicit (for example, by saying *I don't know/understand what you mean*) is more effective in supporting children's communicative understanding and performance, than when the listener interprets the speaker's utterance (for example, by saying *Is it the big one?*). The work by Robinson and Robinson (1982; 1983; see also Lyster, 2004) has shown that metacommunicative moves support both children's communicative understanding and performance. Furthermore, Wegerif et al. (1999) already emphasized the importance of the use of conversational ground rules to support classroom talk. We argue that these conversational ground rules also support talk about talk as they can be used to reflect on children's communicative performance.

1.2. Children's oral communicative competence

From a sociocultural perspective, children's oral communicative abilities are considered an important mediator for self-regulation, learning, and thinking (van Oers, Wardekker, Elbers, & van der Veer, 2008; Whitebread, Mercer, Howe, & Tolmie, 2013). According to Mercer (2008), "learning is mediated through dialogue" (p. 35). Consequently, investing in children's oral communicative competence may contribute to the improvement of the quality of classroom dialogue and children's learning. Furthermore, children's oral language abilities are positively related to their social participation and acceptance (for example, van der Wilt, van Kruistum, van der Veen, & van Oers, in press). As oral language competence is the first language-based competence children develop – and a prerequisite for children's participation in classroom talk – we argue it is important for teachers to pay attention to it from an early age in a systematic and goal-oriented way. Following the work of Schiefelbusch and Pickar (1984) and Celce-Murcia (2008), we use the term oral communicative competence to emphasize that it is about the combination of knowledge, skills and attitudes that "enables a speaker to communicate effectively and appropriately in social contexts" (Schiefelbusch & Pickar, 1984, p. ix; see also; Embrechts, Mugge, & van Bon, 2005). In order for children to communicate successfully, it is not enough only to focus on the systemic and formal aspects of language use (grammar, pronunciation; see Celce-Murcia, 2008; Embrechts et al., 2005), but also to include a more communicative focus which entails practicing the different aspects of social interaction in meaningful activities such as classroom talk (Roth & Spekman, 1984). However, a recent review of empirical research on oral language education in the Netherlands and Flanders indicated that much remains unknown about how teachers can support children's oral communicative abilities through classroom talk (Bonset & Hoogeveen, 2011).

1.3. Current study and research question

In the current study, we investigated the effect of an intervention aimed at increasing the use of productive talk and metacommunication in classroom talk on the development of young children's oral communicative competence. As the relation between dialogically organized classroom talk and children's content learning is not yet firmly established, we also studied the effect on children's subject matter knowledge. Finally, we were interested in the following control variables: sex, age, home language, and socioeconomic status. Previous research has indicated that there might be early sex differences in children's language abilities, with girls performing slightly better than boys (Wallentin, 2008). This sex difference, however, disappears over time. Furthermore, young children's language abilities rapidly increase over time when they enter preschool (for example, Nærland, 2011; Tuijl & Leseman, 2007). A longitudinal study by Tuijl and Leseman (2007) suggests that a non-Dutch home language moderates the growth in language abilities for children attending preschool. Furthermore, a strong relation is often reported between socio-economic status and children's language proficiency and academic achievement (Bradley & Corwyn, 2002).

Using a pre-test-intervention-post-test design, this study was conducted to answer the following research question: What is the effect of an intervention aimed at increasing teachers' use of productive talk and metacommunicative moves on the level of young children's oral communicative competence and subject matter knowledge?

2. Method

2.1. Participants

A total of 21 teachers (and 21 classes) from 11 primary education schools in the Netherlands participated in the study: 12 in the

intervention condition (PCD-condition) and nine in the comparison condition (non-PCD condition). Each participating school contained a maximum of three participating teachers. The schools were distributed throughout the Netherlands and varied in size, population, and context (i.e., urban vs. rural). All participating teachers were women with a mean age of 43.7 years (range 30–58 years). Their average amount of teaching experience was 15.6 years. Table 1 shows descriptive statistics for the most important teacher and classroom characteristics separated for the comparison and the intervention group. Independent-samples *t*-tests showed no significant differences between participating teachers in both conditions for age, years of experience, class size, and number of work days per week. However, a statistically significant difference was found between classrooms for socioeconomic status (class mean), $t(467) = -8.31$, $p < .001$, with children from classrooms in the intervention condition on average having a higher socioeconomic status.

469 children (52.9% boys) participated in this study of whom 273 (58.3%) were in the intervention group and 196 in the comparison group. The children were aged between 3.8 years and 6.5 years, with a mean age of 5.0 years. In primary schools in the Netherlands, it is common practice to have children between roughly 4.0 and 6.5 years of age in one classroom. Table 2 shows descriptive statistics for children's most important demographics. Independent-samples *t*-tests indicated that there were no significant differences between children in both conditions for age, sex, home language, and ethnicity.

2.2. Procedure

In order to recruit teachers to participate in this study, we sent out a call to the schools in our network and via several social media such as Twitter. Interested teachers were visited to give them more information about the study. They were then alternately assigned to either the intervention or comparison group in order of date of visitation (teachers from the first school visited were allocated to the intervention condition, teachers from the second school to the comparison condition, etc.). The study ran during the school year 2014/2015.

Teachers in the intervention condition participated in a Professional Development Program (PDP) on productive classroom dialogue. The PDP consisted of a 2–3 h workshop on productive classroom dialogue, five classroom conversations (once a week) which enabled teachers to practice the use of productive talk and metacommunicative moves, followed by reflection sessions in which episodes from the video recordings of those lessons were watched. In these reflection sessions, there was a strong focus on pedagogy (i.e., productive classroom talk) as this is seen as vital for the effectiveness of PDP's (van Veen, Zwart, Meirink, & Verloop, 2010). During the workshop, the theory of dialogic classroom talk, the talk moves (Appendix A) and the development of children's oral communicative competence were discussed and video examples of classroom talk were watched and analyzed.

All teachers orchestrated the same six classroom conversations related to the theme 'what animal is that?' In each classroom conversation, the children and teacher talked about a different animal and discussed, for example, how this animal eats, whether it is a male or female, if they would want it as a pet and why, etc. In the first classroom conversation children talked about the platypus and in the final conversation they talked about the ladybug. In both conditions, the first and final classroom conversations were video recorded. In this article, we will not report on the analysis of the video observations, but rather focus on the effectiveness of our intervention for the development of children's oral communicative competence and subject matter knowledge.

All children were individually tested on their oral communicative competence and their knowledge about the animals discussed in the different conversations one week before (pre) and one week after (post) the six-week intervention. Tests were administered at a quiet place inside the school by one of the trained test-assistants. Administration of the tests took approximately 25 min per child. All tests were audiotaped so they could be scored afterwards.

Background information on children (age, sex, home language, socioeconomic status) was obtained via the administration offices of each of the participating schools. Furthermore, all participating teachers filled out an online questionnaire in order to obtain background information.

Table 1
Descriptive statistics for teacher and classroom characteristics separated for the comparison and intervention group.

	Comparison group ($N_{non\text{-}PCD} = 9$)			Intervention group ($N_{PCD} = 12$)		
	M	*SD*	Range	*M*	*SD*	Range
Age (years)	43.11	9.79	30–56	44.16	9.62	30–58
Years of experience	14.33	6.44	7–25	16.58	8.62	4–30
Class size	22.22	5.12	13–28	23.17	4.99	14–29
Days working	3.28	1.48	1.0–5.0	3.59	1.00	1.5–5.0
Socioeconomic status	2.39	0.41	1.8–3.0	2.64	0.23	2.2–3.0

Note. Socioeconomic status was aggregated on the classroom level and was measured by averaging both parents' level of education (1 = low, 2 = medium, 3 = high).

2.3. Measures

2.3.1. Oral communicative competence

Children's oral communicative competence was measured using the Nijmegen Test of Pragmatics (NPT; Embrechts et al., 2005). The reliability and validity of the NPT are high (Cronbach's alpha = 0.92; Embrechts et al., 2005). The NPT measures the pragmatic skills of children aged 4–7 and was found to be reliable in the current study (omega = 0.91, GLB = 0.94, Cronbach's alpha = 0.91; based on pretest scores). In order to establish the inter-rater reliability, 5% of all administered tests were independently scored by a second rater. With a Cohen's Kappa of 0.86 the inter-rater reliability of the NPT was found to be strong (Landis & Koch, 1977; McHugh, 2012).

The NPT uses a scale model of a house, nine associated pictures of different rooms in the house, and a standardized protocol. Through a story about the family living in the house, different communicative and conversational responses are elicited (for example, test leader: *"What game do you like to play?";* Child: *[names a game]*; test leader: *"I don't know that game. How is it played?";* child: *[gives an explanation]*; as soon as there is an occasion, the test leader interrupts by saying: *"Huh, I don't understand that"*). 37 items of the NPT were used in this study and each item was scored dichotomously: 1 for a correct and 0 for an incorrect response (in case of the previous example, the child has to show it can repeat something when his/her previous reply is unclear: *[repetition or verbal explanation of what the child said before = 1]* and *[no response or inadequate response = 0]*).

2.3.2. Subject matter knowledge

Children's subject matter knowledge was measured using a verbal Subject Matter Knowledge Test (SMKT) with 15 items that was developed within the context of this study. Children had to answer several questions related to the animals that were discussed in the six different classroom conversations (see Section 2.2). The SMKT contains 12 closed and three more open questions and several pictures of animals that are used to elicit a response (for example, test leader: *"On which of the following four images do you see an ant?"* Child: *[has to point at the right image].* An example of an open question: *"What do you know about ants? Can you tell more?"*

Table 2
Descriptive statistics for children's demographic information for the total sample, comparison group, and intervention group.

	Total sample	Comparison group	Intervention group
N	469	196	273
Age range	3.8–6.5	3.9–6.5	3.8–6.4
M (*SD*)	5.0 (7.7)	4.9 (7.9)	5.0 (7.5)
Sex			
Girls	47.1%	46.9%	47.3%
Boys	52.9%	53.1%	52.7%
Home language			
Dutch as first language	80.8%	83.2%	79.1%
Dutch as second language	19.2%	16.8%	20.9%
Dutch ethnicity	73.8%	73.0%	74.4%

Note. Home language refers to the dominant language spoken at home. Mean age in years, standard deviations in months.

Child: *[has to tell what (s)he knows about ants]*). 12 items were scored dichotomously: 1 for a correct and 0 for an incorrect response. On the three open questions children could score 0, 1 or 2 points depending on the richness of their answer (in case of the previous example: *[the child knows a lot about ants and can tell three or more different things = 2]*, *[the child knows two different things = 1]* and *[the child knows only one thing, talks about something else or tells things that are incorrect = 0]*). Analysis showed this test to have an acceptable internal consistency (omega = 0.71, GLB = 0.78, Cronbach's alpha = 0.69). Furthermore, 5% of all administered tests were independently scored by a second rater to establish the inter-rater reliability. With a Cohen's Kappa of 0.85 the inter-rater reliability of the SMKT was found to be strong (Landis & Koch, 1977; McHugh, 2012).

2.3.3. Intervention fidelity

To evaluate the extent to which our intervention was implemented as planned, we focused on five key elements of intervention fidelity: design, Professional Development Program (PDP), intervention delivery, intervention receipt, and intervention enactment (cf., Smith, Daunic, & Taylor, 2007). The *design* of our intervention and associated PDP were piloted in an earlier study (van der Veen, van Oers, & Michaels, 2014). Classroom conversations were designed to take place over approximately 30 min once a week and were found to fit appropriately within the curriculum. Pre- and post-observations in both conditions indicated that teachers adhered to our teacher manual and were able to conduct all six classroom conversations accordingly.

Our *PDP* was delivered by the first author of this article who had experience in teacher professional development. Teachers in the intervention condition all received the same PDP to ensure systematic delivery across teachers and to maximize fidelity of *intervention delivery*. Furthermore, the first author had weekly meetings with teachers in the intervention condition to reflect on their performance, to answer questions related to the intervention and to ensure they understood the intervention (*intervention receipt*).

To evaluate *intervention enactment*, all video recorded post-observations for teachers in both conditions were viewed and checked against our teacher manual. No abnormalities were found. Furthermore, a 5-min episode of each video was coded for productive talk and metacommunicative moves (Appendix A). In order to capture teacher-student interaction, we started coding 5 min after the start of each classroom conversation. Frequency of productive and metacommunicative teacher moves as percentage of the total number of teacher turns were compared between conditions. An independent-samples *t*-test indicated that teachers in the intervention condition used significantly more productive talk moves (44.12% of teacher turns coded as productive talk moves) compared to teachers in the comparison condition (16.46% of teacher turns coded as productive talk moves), $t(19) = -4.69$, $p < 0.001$.

2.4. Analyses

2.4.1. Missing data

In the final data set, there were multiple missing data points on the outcome variables (i.e., children's oral communicative competence and subject matter knowledge) and pre-test scores. Percentages of missing data points ranged from 7.2 until 7.5% for the outcome variables, and 4.1 until 4.3% for the pre-test scores. Multiple imputation, a strategy aiming to replace missing data (Little & Rubin, 2002) was used after verifying whether missing values were randomly distributed. Little's MCAR test indicated data was not Missing Completely At Random (MCAR), $\chi^2(11) = 33.94$, $p < 0.001$. However, missing data on children's pre- and post-test scores were beyond our control (e.g., children were not at school because of illness, moved to another school or class, etc.) and, therefore, the distribution of missing values is unknown (cf., Schafer & Graham, 2002). We decided to assume our data to be Missing At Random (MAR). Schafer and Graham (2002) argue that this assumption of MAR cannot be tested, but Collins, Schafer, and Kam (2001) demonstrated that a false assumption of MAR hardly influences the estimates and standard errors. Furthermore, Schafer and Graham (2002) argue that in most realistic applications "departures from MAR are not large enough to effectively invalidate the results of an MAR-based analysis" (p. 164). Thus, assuming our data to be MAR, even though they might in fact Not be Missing At Random (NMAR) would not have detrimental effects on our imputed estimates and standard errors. Furthermore, replacing missing values was necessary as analyses required complete and comparable datasets and appropriate sensitivity (power) of the statistical tests. Using the multiple imputation procedure in SPSS 21.0, $m = 50$ complete datasets were produced in which missing data points were imputed with estimated values. The imputed datasets were used in subsequent analyses.

2.4.2. Multilevel analyses

To answer our research question, multilevel modeling was applied in order to correct for the hierarchical structure of the data. Measurements of children's outcome variables (level 1, $N = 469$) were nested within classrooms (level 2, $N = 21$), nested within schools (level 3, $N = 11$). For this purpose, linear mixed model procedures (SPSS 21.0) with maximum likelihood (ML) estimations were carried out following the procedures of Snijders and Bosker (2004). As a first step, intercept-only models (null models) were fitted for both dependent variables, estimating the mean level of children's oral communicative competence and subject matter knowledge, while taking the variance for level 1 (children), level 2 (classrooms), and level 3 (schools) measurements into account. To test the multilevel structure of the data, the amount of variance explained on each level was calculated using intraclass correlation coefficients. The intraclass correlation coefficients (ICC) in our intercept-only models indicated that the school level did not

explain additional variance above the classroom level. Therefore, analyses were continued by fitting random intercept, two-level hierarchical models for each of the outcome variables.

Subsequently, control variables were included in the intercept-only models at both the child and classroom level, while testing the main effects and interaction effects. On the child level, sex (dummy coded, boy = 0), age, children's pre-test scores on respectively the NPT and SMKT, children's home language (dummy coded, Dutch as a first language = 0) were included. On the classroom level, condition (dummy coded, comparison condition = 0) and an aggregated class mean of children's socioeconomic status were added. Children's socioeconomic status was measured by averaging both parents' levels of education (1 = low, 2 = medium, 3 = high). For the sake of parsimony, control variables that yielded non-significant results in all of the model estimations were excluded from the final model. Control variables were included one-by-one using a step-up model building strategy (West, Welch, & Galecki, 2007). According to Kim, Anderson and Keller (2013; see also Ryoo, 2011) "the step-up approach tends to identify the true model in simulations more effectively than the top-down method" (p. 408).

Each model was compared with the intercept-only model using AIC and BIC fit indices (lower indices indicate a better fitting model) and likelihood ratio tests to test for model improvement. Furthermore, decisions on whether parameters should be included in the final model were also made based on the theoretical relevance of each parameter for the final model. Effect sizes were calculated using the procedures suggested by Tymms (2004), using Cohen's *d* for comparing differences between groups and Pearson correlation coefficients (*r*) for comparing the relationship between two continuous variables. All results were tested with an alpha of 0.05.

3. Results

3.1. Examples of productive and non-productive classroom dialogue

For illustrative purposes and to support replication of our study, in Tables 3 and 4 we will provide two examples of post-observation interaction fragments from classroom conversations in the intervention and comparison condition. These examples reflect the previously reported higher frequency of productive talk moves in the intervention condition (see Section 2.3.3). Please note that this does not mean that productive talk moves were absent in classroom conversations in the comparison condition. In both examples, the topic of the classroom conversation is the ladybug. Children have seen a short movie on a flying ladybug and are talking about how the ladybug can fly. Table 3 shows an excerpt from a post-observation classroom conversation in the intervention condition. In this productive classroom dialogue, the teacher uses several talk moves to give children space to say more, listen to one another, reason and think together. Furthermore, children's utterances are relatively long, they take an effort to understand each other, and they are referring to other students' ideas and are thinking with each other without the teacher's mediation.

Table 4 is a fragment from a post-observation classroom conversation in the comparison condition. This excerpt is an example of a non-productive classroom dialogue that is dominantly teacher-steered and in which the teacher uses many Initiation-Response-Evaluation sequences. The teacher asks mainly closed-questions (initiations), followed by a short response by a student, after which the teacher evaluates these responses. Furthermore, in this example students do not think together, address each other or build upon each other's ideas.

3.2. Preliminary analyses

The descriptive statistics of the pre- and post-test measures are presented in Table 5 for the total sample and both conditions separately. Children's oral communicative competence pre-test scores did not differ between conditions, $F(1, 448) = 0.11$, $p = 0.74$. Children's subject matter knowledge pre-test scores, however, did significantly differ between conditions, $F(1, 448) = 24.21$, $p < 0.001$. Children in the intervention condition scored higher on the subject matter knowledge pre-test ($M = 10.05$, $SD = 2.93$) than children in the comparison condition ($M = 8.68$, $SD = 2.89$).

3.3. Multilevel analyses

3.3.1. Intercept-only multilevel models

In Table 6, the intercept-only multilevel models for measures of children's oral communicative competence and subject matter knowledge are presented.

Intraclass correlations indicated that 14.56% of the variance in oral communicative competence was attributable to the classroom level. Furthermore, 14.97% of the variance in measures of children's subject matter knowledge could be attributed to variability on the classroom level. This indicated that the use of multilevel modeling was warranted (Heck, Thomas, & Tabata, 2010; Hox & Roberts, 2010).

3.3.2. Final multilevel models

In this paragraph, we will present the final multilevel models for each of the outcome variables. Table 7 shows the final models for children's oral communicative competence and subject matter knowledge post-test scores. The multilevel analyses revealed there is a statistically significant and strong effect of age on both outcome variables, $t(431) = 2.61$, $p = 0.009$, $r = 0.51$ for children's oral communicative competence and $t(449) = 4.51$, $p < 0.001$, $r = 0.44$ for children's subject matter knowledge with older children having higher oral communicative competence and subject matter knowledge scores. A positive relation with small to moderate effects was found between children's home language and both outcome variables, $t(449) = 2.19$, $p = 0.029$, $d = 0.40$ for children's oral communicative competence and $t(460) = 2.79$, $p = 0.005$, $d = 0.30$ for children's subject matter knowledge. On average, oral communicative competence gain scores (calculated as post-test score−pre-test score) were slightly higher for children with Dutch as a second language ($M = 3.68$, $SE = 0.14$) compared to children with Dutch as a first language ($M = 3.65$, $SE = 0.05$). This difference in gain scores was not statistically significant, $t(2895) = -0.05$, $p = 0.96$. Socioeconomic status (class mean) was found to have a positive but small effect on children's oral communicative competence scores, $t(17) = 1.78$, $p = 0.024$, $r = 0.19$, with high levels of socioeconomic status (averaged on the classroom level) associated with high levels of oral communicative competence. Finally, no significant sex differences were found for both outcome variables.

3.3.3. Effects of condition

The final multilevel model with condition as predictor at the classroom level revealed that the PDP on productive classroom dialogue (i.e., our intervention) had a significant and moderate to large effect on children's oral communicative competence post-test scores as compared to the comparison condition, even after controlling for several variables at the child level and classroom level, $t(18) = -3.44$, $p = 0.001$, $d = 0.62$. The final model with children's subject matter knowledge as outcome variable did not reveal a significant effect of condition, $t(22) = 0.67$, $p = 0.5$, $d = 0.09$. On average, subject matter knowledge gain scores (calculated as post-test score−pre-test score) were higher for children in the comparison condition ($M = 2.53$, $SE = 0.19$) than for children in the intervention condition ($M = 1.77$, $SD = 0.15$).

Table 3
Example of a productive classroom dialogue (post-observation, intervention condition).

01	T:	Dex, will you explain again?	*Say more move (1.2)*
02	S1:	Look, I saw that red part [shield of the ladybug] when it was	
03		going to fly.	
04	S2:	And then – that was actually. I think he means that it is	*S2 refers to*
05		protection for his wing. Not for itself.	*contribution S1*
06	S3:	But also for itself a little bit.	
07	S1:	Yeah, that's also what I meant.	
08	T:	But – um, um – Dana, you were saying that it has protection	*Revoicing move (1.3)*
		for its?	
09	S2:	Um, um – [wings]. For his wings [yeah]. Because otherwise it	
10		will damage its wings and then, then he won't be able to fly	
		very good.	
11	S4:	Like a butterfly [yeah]	
12	S2:	No, because a butterfly has its own wings, but a ladybug goes	*S2 disagrees with S4*
13		like. If you look very carefully – his wings – that red thing	
		goes up like this.	
14	T:	Yeah.	
15	S2:	So that's how they can fly.	
16	T:	So let me see if I understand. You are saying, the red part um	*Revoicing move (1.3)*
17		that's the shield. And that covers his wings [yeah] to protect	
18		its wings? [yeah]. Is that what you're saying?	
19	S2:	Yeah!	
20	S4:	No, its wings are *under* [emphasis added] the red things	*S4 refers to*
		[yeah].	*contribution S2*
21	S1:	Yeah, that's what I meant.	

Note. T = teacher. S = student.

Table 4
Example of a non-productive classroom dialogue (post-observation, comparison condition).

01	T:	What do we see exactly if we look at the picture [of the	*Open question*
02		ladybug] on the screen, what do we see, Almira?	
03	S1:	[inaudible] his wings.	
04	T:	You can see his wings. And what do you see on its wings?	*Initiation (I)*
05	S1:	Dots.	*Response (R)*
06	T:	Yeah, dots, that's correct. And how many dots do you see?	*Evaluation (E) – Initiation (I)*
07	S2:	One, two, three, four, five.	*Response (R)*
08	T:	We can see five dots now, right [stands up and points at the	*Evaluation (E)*
09		picture on the screen]. There are one, two, three dots on this	
10		side, one in the middle, and one here. Could there also be two	*Initiation (I)*
11		dots on the other side?	
12	S3:	Yeah.	*Response (R)*
13	T:	Just like here, right. We just saw that in the movie, huh.	*Evaluation (E)*
14		There were seven dots. There are many different types of	
15		ladybugs. What does the ladybug look like? We have already	
16		mentioned that his back is red. Does any of you know how	*Initiation (I)*
17		that's called? How do we call the back of a ladybug? They just	
18		mentioned it [in the movie].	
19	S4:	A shield.	*Response (R)*
20	T:	It's called a shield, very good Luke, you did pay attention.	*Evaluation (E)*
21		Actually, it has two shields on his back. He can open this side	
22		and it can also open that side. And what is under it, under the	*Initiation (I)*
23		shield, Anton?	
24	S5:	Wings.	*Response (R)*
25	T:	Wings. And are these wings also hard, Sam? Are they just as	*Evaluation (E)*
26		hard as the shield on top, the wings? [Sam nods yes]. You	*Agree-disagree move*
27		think they are. Who thinks something else, Fiona?	*(4.1)*
28	T:	His wings, are they also thick, just as thick as its shield or are	*Initiation (I)*
29		his wings very thin?	
30	S6:	Very thin.	*Response (R)*

Note. T = teacher. S = student.

4. Discussion

This study investigated the effect of an intervention aimed at increasing the use of productive talk and metacommunication in classroom talk on the development of young children's oral communicative competence and subject matter knowledge. Multilevel models revealed that teachers' participation in a Professional Development Program (PDP) on productive classroom dialogue (i.e., the intervention condition) had a significant and moderate to large effect on the level of young children's oral communicative competence. Final multilevel models did not reveal a significant effect of the PDP on the level of children's subject matter knowledge. Our findings add to previous studies on the value of dialogically organized classroom talk for children's (content) learning and development (for example, Kiemer et al., 2015; Nystrand & Gamoran, 1991; O'Connor, Michaels, & Chapin, under review) and their oral communicative competence (van der Veen et al., 2014).

Howe and Abedin (2013) state in their review on classroom dialogue that dialogically organized classroom talk in small-group settings is beneficial for children's learning. They hypothesize

Table 5
Descriptive statistics for children's oral communicative competence and subject matter knowledge pre- and post-test scores per condition and for the total sample.

Measure	Total			Intervention condition			Comparison condition		
	Valid *n*	*M*	*SD*	Valid *n*	*M*	*SD*	Valid *n*	*M*	*SD*
OCC									
Pre-test	450	23.91	8.08	263	24.01	7.83	187	23.76	8.45
Post-test	435	27.68	7.43	253	28.85	6.82	182	26.05	7.94
Knowledge									
Pre-test	449	9.48	2.99	263	10.05	2.93	186	8.68	2.89
Post-test	434	11.62	3.14	252	11.91	3.07	182	11.22	3.21

Note. OCC = Oral communicative competence. $N = 469$, valid *n* reports the number of valid cases prior to multiple imputation of missing data.

Table 6
Intercept-only multilevel models for children's oral communicative competence and subject matter knowledge scores.

Parameter	Oral communicative competence			Subject matter knowledge		
	Estimate	*SE*	*p*	Estimate	*SE*	*p*
Fixed effects						
Intercept	27.20	0.72	***	11.46	0.31	***
Random effects						
Level 2 variance – classroom (τ^2)	8.62	3.62	*	1.55	0.63	*
Level 1 variance – child (σ^2)	50.57	3.51	***	8.79	0.62	***
ICC	0.15			0.15		
Model fit statistics						
−2 log likelihood	3201.44			2383.22		
AIC	3205.44			2387.22		
BIC	3213.73			2395.52		

Note. For effects: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

that this is likely to be the case for whole-class settings, but due to limited availability of quantitative, quasi-experimental studies it remains unknown to what extent dialogically organized talk in whole-class settings is indeed beneficial for student performance (Howe & Abedin, 2013). Orchestrating productive classroom dialogue in whole-group settings is a complex endeavor in which teachers have to manage many things at the same time (O'Connor et al., under review). This is even more complex in early childhood education settings: young children have a hard time regulating their behavior (for example, waiting to take a turn to talk) and are still learning to sit still, concentrate and listen to one another. In our study we were able to confirm Howe's and Abedin's hypothesis and show (1) that teachers in early childhood classrooms are able to learn to orchestrate whole-class dialogic talk through participation in a PDP and using a set of productive and metacommunicative talk moves, and (2) that dialogically organized talk in whole-group settings is beneficial for student performance, in particular children's oral communicative competence.

Apart from the effect of our intervention for children's oral communicative competence, our analyses showed that age and children's dominant home language had a strong effect on children's oral communicative competence scores. The effect of age is in line with previous studies and indicates that children's language and communicative abilities rapidly develop during early childhood education (for example, Nærland, 2011) and give them increasing possibilities to successfully interact in different social situations. Our multilevel model indicates that participating in dialogically organized classroom talk contributes to the development of children's oral communicative competence even after controlling for age. Furthermore, our analyses show that participating in productive classroom dialogue is equally beneficial for children with Dutch as their first language spoken at home as for children with Dutch as their second language; children with Dutch as a second language even benefit slightly more. This is in line with a previous study by Tuijl and Leseman (2007) in which they showed that children with a non-Dutch home language gained higher verbal intelligence scores from attending preschool in which there was a strong focus on language education. In dialogic classroom talk, children are supported to share ideas, listen to one another, reason, and think together. As such, children are given the

Table 7
Final multilevel models for children's oral communicative competence and subject matter knowledge scores.

Parameter	Oral communicative competence			Subject matter knowledge		
	Estimate	*SE*	*p*	Estimate	*SE*	*p*
Fixed effects						
Intercept	1.38	2.81		0.40	0.96	
Child level						
Sex (0 = boy)	−0.58	0.40		0.26	0.21	
Age (pre-test)	0.08	0.03	**	0.07	0.02	***
Pre-test	0.70	0.03	***	0.61	0.04	***
Home language (0 = Dutch as first language)	1.19	0.54	*	0.83	0.30	**
Classroom level						
Condition (0 = Comparison)	−1.83	0.53	***	0.25	0.40	
Socioeconomic status (class mean)	1.80	0.79	*	–	–	
Random effects						
Level 2 variance – classroom (τ^2)	0.40	0.41		0.44	0.33	*
Level 1 variance – child (σ^2)	16.65	1.21	***	4.61	0.21	***
ICC		0.02			0.09	
Model fit statistics						
−2 log likelihood	2613.39			2035.75		
AIC	2631.39			2051.75		
BIC	2668.59			2084.82		
Likelihood ratio test	$\chi^2(6) = 588.05$		***	$\chi^2(5) = 347.47$		***

Note. For effects: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Dashed cells indicate parameters that were not estimated in the final model for the sake of parsimony (estimates were not significant in any of the previous models).

possibility to meaningfully practice the different aspects of communication and social interaction that will, consequently, support the development of their oral communicative competence.

Surprisingly, our analyses did not reveal a significant effect of condition for children's subject matter knowledge. Gain scores even indicated that children gained more subject matter knowledge from participating in non-productive classroom dialogue compared to participating in dialogically organized classroom talk. Previous research, however, suggests that participating in dialogically organized classroom talk is positively related to children's performance in English literature (Nystrand & Gamoran, 1991), mathematics (O'Connor et al., under review), and non-verbal reasoning (Wegerif et al., 1999). How can the results of our study be explained? A first explanation might be that we used a verbal knowledge test that was not standardized or piloted due to time constraints. A second explanation is related to the equivalence of content (or subject matter) between both conditions in our study. Although all teachers carried out the same six classroom conversations on animals and had access to the same teaching materials, we were not able to control the ideas shared and questions asked between both conditions. Our video observations indicate that students in the intervention condition often came up with unexpected questions or ideas that deepened their reasoning (see also O'Connor, 2001; O'Connor et al., under review; Table 3) and guided the conversations in new directions. Consequently, classroom conversations in the intervention condition did not always cover all the content from the teaching manual. Conversations in the comparison condition were often highly structured, teacher-steered, and therefore covered mostly all the content from the teaching manual (see Table 4). A difference in content or subject matter knowledge between both conditions might explain why children in the comparison condition gained more subject matter knowledge over the duration of our intervention. In order to gain more insight into the value of dialogically organized classroom talk for children's learning and development, future research should also aim to control content between conditions and classrooms as much as possible. Another limitation of our study is that we measured oral communicative competence – a complex, multifaceted competence – using only one instrument. Future studies should attempt to take more elements of children's oral communicative competence into account, such as listening skills, attitude and theory of mind.

To conclude, the present study showed that an intervention in which teachers learn to use productive talk moves and metacommunicative moves in the context of whole-group classroom talk with young children has a significant and moderate to large effect on the development of children's oral communicative competence. Future research should further explore whether this effect sustains over time and how each of the talk moves and associated teacher-child interaction sequences contribute to this effect. Furthermore, it is worthwhile to explore the differences in children's learning, thinking and communicative development in the context of productive and non-productive classroom dialogue. For example, do children who are inducted into dialogically organized classroom talk over a longer period of time learn to think differently (i.e., reasoning, thinking together, listening) and do they become more competent communicators? The results of those studies and analyses might add to the available empirical evidence of the benefits of productive classroom talk, and support teachers in establishing a dialogically organized classroom culture. Howe and Abedin (2013) suggested "it is time to take risks" (p. 346) and to also use quantitative methods to examine whether dialogically organized classroom talk is more beneficial than other modes of organization. Was it worthwhile to take this risk? We believe it was.

Acknowledgements

The authors would like to thank all teachers and children who participated in the 'MODEL2TALK' project and the test-assistants for helping with the collection of the data.

Appendix A. Descriptions and examples of the productive talk and metacommunicative moves.

Teacher talk moves	Description	Example
1. Share, expand, clarify	**Encourage children to share, expand and clarify their initial ideas or utterances**	
1.1 Time to think	Verbally encourage children to take some time to think	*"Let's all take a minute to think about it"*
1.2 Say more	Encourage children to say more about their initial idea or press them to clarify their thoughts	*"Can you say more about what you mean with shield?"*
1.3 Revoicing	Rephrasing or restating (parts) of a child's utterance, in order to verify, rebroadcast or position the initial utterance and to create a space for the child to agree/disagree, say more, negotiate, etc.	*"So you are saying that the red part, that's his shield?"*
2. Listen to one another	**Encourage children to listen to one another**	
2.1 Repeat or rephrase	Encourage children to repeat or rephrase another's child contribution in order to stimulate children to listen to one another and taking other's ideas seriously	*"Who can repeat what Dana just said?"*
3. Reasoning	**Ask children for evidence and stimulate them to deepen their reasoning**	
3.1 Why	Press children for reasoning or ask them to come up with the rationale behind their initial idea	*"Why do you think the ladybug needs a shield?"*
3.2 Challenge or counterexample	Encourage children to come up with a counterexample or challenge an initial claim	*"Does it always work that way?"*
4. Think with others	**Encourage children to think with each other and build on each other's ideas**	
4.1 Agree or disagree	Encourage other children to agree or disagree with one child's initial idea	*"Jurre, do you agree with Sara's idea? Why?"*
4.2 Add on	Encourage children to add on or respond to someone else's idea	*"Who can add on Sanne's idea about the shield of the ladybug?"*
4.3 Explaining someone else	Ask children to explain what someone else means	*"Who can explain what Douwe means when he says that?"*
5. Metacommunication	**Encourage children to reflect on their communicative performance and the understandability of their oral messages**	
5.1 Metacognitive guidance	Explicitly indicating a problem of non-comprehension and encourage children to rephrase their initial message	*"I don't know/understand what you mean"*
5.2 Conversational ground rules	Explicitly pointing at conversational ground rules that apply to the group, negotiate new rules and stimulate children to reflect on their communicative performance	*"Why is it important that we listen carefully to one another?" "What talk rules did we agree upon?"*

References

Bonset, H., & Hoogeveen, M. (2011). *Mondelinge taalvaardigheid in het basisonderwijs. Een inventarisatie van empirisch onderzoek* [*Oral language skills in primary education schools. A review of empirical research*]. Enschede, the Netherlands: SLO.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371–399. http://dx.doi.org/10.1146/annurev.psych.53.100901.135233.

Cazden, C. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.

Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. A. Soler, & P. S. Jordà (Eds.), *Intercultural language use and language learning* (pp. 41–57). Dordrecht, the Netherlands: Springer.

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing-data procedures. *Psychological Methods, 6*, 330–351. http://dx.doi.org/10.1037/1082-989X.6.4.330.

Embrechts, M., Mugge, A., & van Bon, W. (2005). *Nijmeegse Pragmatiek test. Handleiding* [*Nijmegen test for pragmatics. Manual*]. Amsterdam, the Netherlands: Harcourt Test Publishers.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2010). *Multilevel and longitudinal modelling with IBM SPSS*. New York, NY/London, England: Routledge.

Howe, C., & Abedin, M. (2013). Classroom dialogue: a systematic review across four decades of research. *Cambridge Journal of Education, 43*(3), 325–356. http://dx.doi.org/10.1080/0305764X.2013.786024.

Hox, J. J., & Roberts, J. K. (Eds.). (2010). *Handbook of advanced multilevel analysis*. New York, NY: Routledge Academic.

Kiemer, K., Gröschner, A., Pehmer, A. K., & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction, 35*, 94–103. http://dx.doi.org/10.1016/j.learninstruc.2014.10.003.

Kim, J. S., Anderson, C. J., & Keller, B. (2013). Multilevel analysis of assessment data. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment* (pp. 389–424). Boca Raton, FL: CRC Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. http://dx.doi.org/10.2307/2529310.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Hoboken, NJ: Wiley.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition, 26*(3), 399–432. http://dx.doi.org/10.1017/S0272263104263021.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica, 22*(3), 276–282. http://dx.doi.org/10.11613/BM.2012.031.

Mercer, N. (2008). The seeds of time: why classroom dialogue needs a temporal analysis. *Journal of the Learning Sciences, 17*(1), 33–59. http://dx.doi.org/10.1080/10508400701793182.

Michaels, S., & O'Connor, C. (2012). *Talk science primer*. Cambridge, MA: TERC.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: professional development approaches for academically productive discussion. In L. B. Resnick, C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through talk and dialogue* (pp. 347–362). Washington, DC: AERA.

Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English, 25*(3), 261–290. Retrieved from http://www.jstor.org/stable/40171413.

Nystrand, M., Wu, L. L., Gamoron, A., Zeiser, S., & Long, D. A. (2003). Questions in time: investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes, 35*(2), 135–198. http://dx.doi.org/10.1207/S15326950DP3502_3.

Nærland, T. (2011). Language competence and social focus among preschool children. *Early Child Development and Care, 181*(5), 599–612. http://dx.doi.org/10.1080/03004431003665780.

van Oers, B., Wardekker, W., Elbers, E., & van der Veer, R. (Eds.). (2008). *The transformation of learning. Advances in cultural-historical activity theory*. Cambridge, England: Cambridge University Press.

O'Connor, M. C. (2001). "Can any fraction be turned into a decimal?" A case study of a mathematical group discussion. *Educational Studies in Mathematics, 46*(1), 143–185. http://dx.doi.org/10.1023/A:1014041308444.

O'Connor, M. C., Michaels, S., & Chapin, S. (2015). "Scaling down" to explore the role of talk in learning: from district intervention to controlled classroom study. In L. B. Resnick, C. Asterhan, & S. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 111–126). Washington, DC: AERA.

O'Connor, C., Michaels, S., Chapin, S. The silent and the vocal: participation and learning in whole- class discussion, Learning and Instruction (under review).

Resnick, L. B., Asterhan, C., & Clarke, S. (Eds.). (2015). *Socializing intelligence through academic talk and dialogue*. Washington, DC: AERA.

Robinson, E. J., & Robinson, W. P. (1982). The advancement of children's verbal referential communication skills: the role of metacognitive guidance. *International Journal of Behavioral Development, 5*(3), 329–355. http://dx.doi.org/10.1177/016502548200500304.

Robinson, E. J., & Robinson, W. P. (1983). Communication and metacommunication: quality of children's instructions in relation to judgments about the adequacy of instructions and the locus of responsibility for communication failure. *Journal of Experimental Child Psychology, 36*(2), 305–320. http://dx.doi.org/10.1016/0022-0965(83)90035-8.

Roth, F. P., & Spekman, N. J. (1984). Assessing the pragmatic abilities of children: Part 1. Organizational framework and assessment parameters. *Journal of Speech and Hearing Disorders, 49*(1), 2–11.

Ryoo, J. H. (2011). Model selection with the linear mixed model for longitudinal data. *Multivariate Behavioral Research, 46*(4), 598–624. http://dx.doi.org/10.1080/00273171.2011.589264.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods, 7*(2), 147–177. http://dx.doi.org/10.1037//1082-989X.7.2.147.

Schiefelbusch, R. L., & Pickar, J. (Eds.). (1984). *The acquisition of communicative competence*. Baltimore, MD: University Park Press.

Smith, S. W., Daunic, A. P., & Taylor, G. G. (2007). Treatment fidelity in applied educational research: expanding the adoption and application of measures to ensure evidence-based practice. *Education and Treatment of Children, 30*(4), 121–134.

Snijders, T. A. B., & Bosker, R. J. (2004). *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. London, England: Sage.

Tuijl, C., & Leseman, P. M. (2007). Increases in the verbal and fluid cognitive abilities of disadvantaged children attending preschool in the Netherlands. *Early Childhood Research Quarterly, 22*(2), 188–203. http://dx.doi.org/10.1016/j.ecresq.2007.02.002.

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen, & K. Elliot (Eds.), *But what does it mean? Calculating effect sizes in educational research* (pp. 55–66). Slough, England: National Foundation for Educational Research.

van der Veen, C., van Kruistum, C., & Michaels, S. (2015). Productive classroom dialogue as an activity of shared thinking and communicating: A commentary on Marsal. *Mind, Culture and Activity: An International Journal, 22*(4), 320–325. http://dx.doi.org/10.1080/10749039.2015.1071398.

van der Veen, C., van Oers, B., & Michaels, S. (2014, September). Promoting productive classroom dialogue in early childhood education. In *Paper presented at the fourth ISCAR Conference. Sydney, Australia*.

van Veen, K., Zwart, R., Meirink, J., & Verloop, N. (2010). *Professionele ontwikkeling van leraren. Een review studie naar effectieve kenmerken van professionaliseringsinterventies van leraren* [*Teacher professional development. A review study into the effective characteristics of interventions for teacher professional development*]. Leiden, the Netherlands: ICLON.

Wallentin, M. (2008). Putative sex differences in verbal abilities and language cortex: a critical review. *Brain & Language, 108*(3), 175–183. http://dx.doi.org/10.1016/j.bandl.2008.07.001.

Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: an empirical investigation of a possible sociocultural model of cognitive development. *Learning and Instruction, 9*(6), 493–516. http://dx.doi.org/10.1016/S0959-4752(99)00013-4.

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. New York, NY: Chapman Hall/CRC.

British Journal of Educational Psychology, Monograph Series II: No. 10. In Whitebread, D., Mercer, N., Howe, C., & Tolmie, A. (Eds.). (2013), *Self-regulation and dialogue in primary classrooms*. Leicester, England: BPS.

van der Wilt, F., van Kruistum, C., van der Veen, C., & van Oers, B. (2015). Gender differences in the relationship between oral communicative competence and peer rejection: an explorative study in early childhood education. *European Early Childhood Education Research Journal, 25*(2). http://dx.doi.org/10.1080/1350293X.2015.1073507 (in press).