Learning outcomes of exploratory talk in collaborative activities

Proefschrift voorgelegd tot het behalen van de graad van doctor in de onderwijswetenschappen aan de Universiteit Antwerpen

Jan T'Sas

Promotoren
Prof. dr. Mathea Simons
Prof. dr. Sven De Maeyer

Faculteit Sociale Wetenschappen
Antwerp School of Education
Antwerpen 2018

Universiteit Antwerpen

Universiteit Antwerpen

Faculteit Sociale wetenschappen

Antwerp School of Education

Learning outcomes of exploratory talk in collaborative activities

Proefschrift ingediend tot het behalen van de graad van doctor in de onderwijswetenschappen aan de Universiteit Antwerpen door

Jan T'Sas

Antwerpen 2018

Promotors:

Prof. dr. Mathea Simons

Prof. dr. Sven De Maeyer

Members Doctoral Commission:

Prof. dr. Rita Rymenans

Em. prof. dr. Frans Daems

Prof. dr. Tom Smits

Members Doctoral Jury:

Prof. dr. Tom Smits (chairman), University of Antwerp

Prof. dr. Rupert Wegerif, University of Cambridge (UK)

Em. prof. dr. Roger Standaert, University of Ghent

Dutch Title:

Leereffecten van exploratieve gesprekken tijdens groepswerk

D/2018/12.293/12 ISBN: 9789057285844

To my mother and father, long gone but never out of my heart and mind.

Table of contents

Acknowledgements

A few years ago Flemish radio Klara - all culture and classical music - invited famous Belgian artists to tell their listeners which word of our language they did not like. At the same time, all Klara listeners were challenged to find an alternative which the artists would find more beautiful. And so it happened that poet and novelist Bart Moeyaert launched his call to find another word for 'applause' (which he was dying to get rid off). Lots of alternatives came, until Moeyaert chose his favourite and that was 'clappreciation'. To me, too, the word sounded just right, and so, thanking Bart Moeyaert and radio Klara, I would like to express my loudest clappreciation to all those I owe this dissertation to.

First I want to clap hands for prof. Neil Mercer, who came down from Cambridge to give a lecture about exploratory talk at Antwerp University in November 2010. His story was completely unknown to me, but after his lecture I realised I had found the missing link in the didactic approach of speaking and listening skills, which I had been teaching for some years. A few years later, when I mailed him I would try and replicate his research in my own country, he immediately invited me to come over to Cambridge and discuss it. I am still grateful for his and his colleague Lynn Dawes' hospitality and good advice.

I also want to thank Neil's colleague, Rupert Wegerif, for being part of my doctoral jury. As soon as I started reading about exploratory talk, Rupert's name turned up nearly as often as Neil's, making clear that I had encountered another pioneer I would want to meet. As Rupert is deepening his theoretical insights in dialogue for learning I will be glad to keep following his scientific and philosophical tracks. Likewise, I want to thank Roger Standaert for the inspiring talks we had throughout the years we shared in the world of education and for his warm support of my research.

My next round of - very warm - clappreciation goes to my promotors, Mathea Simons and Sven De Maeyer, who - I admit - had some trouble convincing me to start this research project, but as soon as they had lit the research fire inside me, they knew very well how to keep it burning. I am forever in their debt for their never lacking enthousiasm, continuing advice and support, and their critical, constructive thinking.

A third round of clappreciation goes to the doctoral committee, Tom Smits, Rita Rymenans and Frans Daems, as well as Mathea Simons and Sven De Maeyer. Their vast experience proved more than valuable to keep me on track. I also thank the members of the doctoral jury for hearing me out and granting me this phD-title: Rupert Wegerif, Roger Standaert, Tom Smits, Mathea Simons and Sven De Maeyer.

Fourth, special thanks to Ria Van den Eynde, my temporary colleague at Karel de Grote Hogeschool which funded the main part of this research. Ria 'popped in' as the experiments were already up and running, but her contribution I consider beyond value.

I very much thank the teachers of the research schools and their pupils for walking this path all the way with great zest and dedication. These are Miss Kelly, Miss Marijke, Miss Hilde, Miss Daisy, Master Kurt, Miss Chloë, Miss Kathleen, Miss Carmen, Miss Stéphanie, Miss Ameryllis en Miss Glennys of De Zeppelin Haasdonk, GBS Beveren, Immaculata Deurne, Sint-Eduardus Schoten en Vincent Ekeren. During the experiments, when I was not frantically walking from one camera to another to see if all was functioning well, we shared lots of short and interesting reflections. My special thanks goes to Miss Chloë who also chose my dataset as a research ground for her own master thesis and added new insights for follow-up research. Further, my sincere gratitude goes to Eva Soenens, Annelies Verküringen en Naomi Peeters, all students of Antwerp University, for their thorough transcription of more than 60 hours of conversation, and to my students Helena Lemmens and Bert Van Poeck for their final reading of this dissertation.

Fifth, finding myself on the brink of getting sore hands, I thank my university colleagues, especially Rudi, Jordi and Marc, whose curiosity in my proceedings always gave me a stress-killing good feeling. Furthermore, I will always be grateful for the professional opportunities Antwerp University has given me, including the positive vibes of many colleagues as soon as they heard ‘what was going on’.

I am also very glad I had the chance to share this intellectual adventure with my dearest wife, Annick. Not only did we have fruitful conversations about various research issues, she was always there to support and encourage me, saying things like ‘You will only get a glass of wine after you have finished the next chapter’. Moreover, she has accomplished something few people have been able to do so far, which is having me say ‘no’ to other interesting projects and initiatives as long as I was working on my doctoral dissertation. I also thank my stepdaughter Mieke who from the beginning showed a lot of interest in exploratory talk and shared her valuable experiences as a kindergarten teacher with me.

Finally, to my children Alexander, Hannah and Joachim, all in their (early) twenties as I am writing this: it must have something to do with the genes, but do not worry if things take more time than you think, for we are a family of late bloomers. When I had your age and was a student myself, I would never ever have thought I would acquire a phD forty years later. True, it takes a master’s degree to start with, but it also requires inspiration and transpiration, professional enthousiasm, a continuous interest in things and in people, the art of connecting and seeing links, positive thinking and an attitude of sincere modesty, no matter how good an expert you have become in your field. Criticism and curiosity have always been valuable teachers to me and I am grateful to have met smart people who not only inspired me but who included me in various educational projects, just because they saw an added value in doing so. In short, when talent shows, it is more likely to find fertile soil, but at the same time, it takes an open eye to find that soil (Okay, okay, I will stop preaching now).

To all of you, my most sincere clappreciation!

Introduction

In November 2010 Antwerp University, the university I have been working for twenty years now, invited prof. Neil Mercer from the University of Cambridge to give a lecture on 'Language and education: improving pupils skills in communicating, learning and reasoning'. At that time Neil Mercer, who is now emeritus, was professor of Language and Communications at the Open University of Cambridge, where he was also director of the Centre for Language and Communications and director of the Centre for Research in Education and Educational Technology. The title of his lecture intrigued me. For twelve years I had been working part time as a teacher educator. My domain is Dutch didactics and among my specialised subjects are speaking and listening skills, and how you teach them at school. Research on those skills was (and is) very limited in Flanders and The Netherlands, so when Neil's lecture was announced, I knew I would welcome any new ideas to improve my teaching practice.

It was during that lecture, which lasted only 45 minutes, that I first heard about exploratory talk and its educational potential. I was thoroughly impressed. Looking back on that evening I keep wondering why Mercer's typology of talk was completely unknown to me and - as I soon found out - to most of my colleagues. But that evening it felt as if I had discovered a missing link in the didactic approach of speaking and listening skills. Exploratory talk emphasises the strong relationship between language and learning in classroom activities and makes it very concrete. Hearing that, I witnessed Vygotsky's 'old' theories jump to life again.

The day after his lecture, Neil Mercer and Lynn Dawes, with whom he shares a number of studies and publications, gave a workshop on the topic. I convinced some of my students to join me and the whole morning we did varied classroom excercises while practicising exploratory talk and reflecting on ways to implement it in education. At that time I already was a regular trainer at the Centrum Nascholing Onderwijs, Antwerp University's in-service training institute for teachers and educators. Soon after Mercer's lecture I developed a training on group work and included some of his ideas. I wrote a book on group work and devoted a large chapter to exploratory talk, including lots of concrete lesson materials which I had either developed or translated. Then already it struck me that few school books promoted dialogic teaching[1], let alone exploratory talk. As I expected, Flemish teachers were equally 'ignorant' but also pleasantly surprised when we did 'exploratory exercises' during training and the book itself was received as 'an innovative source of inspiration'.

---

[1] 'Dialogic teaching' is a specific educational approach in which teachers and pupils work together to construct knowledge through interaction. Teachers and pupils address learning tasks together, share ideas and consider alternative viewpoints, feel free to voice their opinion without fear of giving 'wrong answers', help each other to reach common understandings and build on their own and each other's ideas. To realise this, 'teachers plan and steer classroom talk based on specific educational goals' (Alexander in Mercer and Hodgkinson, 2008, p. 105; Lyle, 2008).

.

Three years later, after talking about the topic for the *n*-th time, thereby regurarily expressing the regret that exploratory talk and learning was such uncharted territory in Flemish education, some of my colleagues convinced me to start research myself. When I mailed Neil Mercer I wanted to replicate his research in my own country, he immediately invited me to come over to Cambridge and discuss it. Meanwhile, I was happy to have this research partly funded by Karel de Grote Hogeschool. To cut a long story short: this dissertation is the result of seven, nearly eight years of pondering, reading, studying, trying out, experimenting, talking and reflecting.

The dissertation you are about to read is composed of six chapters. After this introduction, Chapter One focuses on the problem. From a broad educational context the focus moves to the actual research topic: constructivist thinking in education, language as a means to construct language, research on speaking and listening skills in Flanders and the Netherlands, the use of language as a learning tool in the classroom, exploratory talk as 'a missing novelty' in both Flemish research and Flemish schools. After that our research questions are formulated, followed by an outline and description of our research design.

Chapter Two comprises the theoretical framework which underpins the studies. Based on a general literature study the role of language for learning is explored on three levels, from a wide to a more specific perspective: (1) learning theories and the role of language, (2) the use of language for learning in the classroom and (3) language and learning in group work.

Chapter Three describes a narrative review that resulted from a standardised search into the concept of exploratory talk in international research databases (RQ 1). In the first part of this review the concept of exploratory talk is defined. To that purpose already existing definitions were identified, (near) synonyms and antonyms were analysed, and the growth of the concept and its contextual use were charted. Next, the review resulted in an overview of the measurement of exploratory talk, answering questions like which methods are used and what variables are taken into account. Finally, the effects of exploratory talk in education are described.

In Chapter Four the design and results are presented of an interventional quasi-experiment through a 12-week training of exploratory talk in five Antwerp primary schools. The research design of this study is outlined as well as the results of a preliminary pilot study, followed by the results of the main study which reports on the use of exploratory talk before (RQ 2), after this 12-week training (RQ 3) and the effects on the problem solving skills of the participants (RQ 4). These results are triangulated with individual characteristics (RQ 5).

In Chapter Five conclusions are drawn about the context of this study, the methodology and the results.

Chapter Six terminates this dissertation. Some issues for discussion are brought up and suggestions are made for further research.

Chapter 1: Problem, research questions and research design

This chapter focuses on the role of language and learning in Flemish education and the problem this study wishes to address. We will argue that, inspired by social constructivist theory, Flemish and Dutch researchers and educationalists studied and acknowledged the importance of language as a learning tool as early as the 1970s. However, evidence of research in the 1990s will show that despite some very promising pioneer projects in schools wide acceptance of this idea stayed out. We will also argue that prior attention for speaking and listening skills has gradually decreased in classroom practice and in educational research, especially when looking at Dutch as a school subject. We will conclude that all this explains why the concept of exploratory talk has never really found its way to our curricula or to classroom practice. It also explains why certain deficiencies in Flemish pupils' oral skills show in national assessment studies.

1. Problem

21st century competencies and constructivist thinking in education

Technology has changed the way we live, work and communicate thoroughly. It demands new skills from us, replacing old ones, and necessitates new approaches in education. At the end of last century Flemish employees were found to score high on skills and attitudes like loyalty, obedience, efficiency, logical thinking and professional ethos. At the same time, they scored below the European average on '21th century skills' such as responsibility, tolerance, communicative skills, creativity, critical thinking, flexibility, working in teams and problem solving (Cincinnato & De Meyer, 2014; Vlaams Parlement, 1998). There is a general agreement that the knowledge, skills and attitudes we needed some thirty years ago no longer suffice for today's requirements (Van den Branden, 2015).

Therefore, which competencies do (young) people need to function effectively as active and responsible citizens, employees and learners in 21st century society and what is the role of language in all this? The answer to these questions may be found on a Europan level, where 'Key Competencies for the 21st century' have been developed. These competencies focus heavily on the skilfull processing and application of knowledge, language and information and of 'learning power' (Van den Branden & Van Gorp, 2000). From an educational point of view, the Key Competencies are not entirely new. They can be traced back to principles of constructivist thinking about knowledge and learning. The principles of constructivism find their origin in the theories of Bruner, Piaget and Vygotksy (Valcke, 2010; Piaget, 1970, 1995; Piaget & Elkind, 1968; Vygotsky, 1978) and were fully implemented in the Flemish educational curricula in the 1990s. In these curricula De Corte's (1996) description of constructivist learning is commonly referred to. He describes it as a constructive, cumulative, self-regulatory, intentional, context bound, collaborative[2] and individual process of acquiring knowledge, giving meaning and developing skills.

Cooperation almost inevitably involves language. In its most effective form it stimulates learning from one another and enhances knowledge construction (Kagan, 2014; Mercer, 1995; van der Aalsvoort & van der Leeuw, 1982). According to Vygotsky (1978) it also facilitates individual learning. From these points of view, there is a reciprocal connection between language and learning. But how and to what extent does this show in education?

In an attempt to pinpoint the problems Flemish education is facing when it comes to consciously and systematically connecting language with learning and to clarify the need for the present study, we will zoom in on the importance as well as the study and classroom practice of listening and speaking skills in education in the Low Countries.

Language as a means to construct knowledge

From a constructivist point of view, language is of key importance for learning, as van der Aalsvoort and van der Leeuw (1992) concluded after an international literature study in the late 1970s and early 1980s. Inspired by the influential British educationalist Douglas Barnes (1976) they promoted language as a means to learn and construct knowledge rather than reducing it to a mere vehicle for the transfer of knowledge, which at that time and long before seemed to be common practice in many classrooms. They strongly recommended education to implement the *interpretative model* (Barnes, 1976) in which knowledge is developed from and together with the pupils, using language. This model, van der Aalsvoort and van der Leeuw (1992) argue, approximates the way we learn in everyday life. In the interpretative model, knowledge is not an object waiting out there to be manipulated one way or another; it is a human construction. Learning itself is a subjective experience of learners. The teacher[3], then, is not a dominant transferer of knowledge, but a learning supervisor who facilitates learning, who guides learning processes and who manages classroom talk. His role is to create an environment that allows and encourages (social) learning. It is not clear to him what exactly his pupils will learn, because

---

[2] When classroom activities in which pupils work together are described, the terms collaborative learning and cooperative learning are often used interchangeably. Both have many features in common, but there are some differences. With collaborative learning, students make individual progress while working together with others and are scaffolded in doing so by the teacher. Further, according to Rockwood III (1995) collaborative learning is connected to the social constructionist's view that knowledge is a social construct. Cooperative learning involves more inherent interdependence without the direct support of the teacher, and greater accountability.
In Flemish education the word 'coöperatief' (cooperative) is used when referring specifically to the aspect of interdependence during collaborative activities (Paelman, 2004). Though the difference between both terms is indeed rather subtle, we will use the term cooperative only when it is used to stress the above mentioned interdepence or when citing authors who use the term explicitly.
[3] Though a teacher can be a man or a woman, for stylistic reasons we will refer to the teacher as 'he' in this dissertation.

each individual learns in a different way. For this reason, he has to familiarise himself with the learning process of his pupils by using active teaching methods. Consequently, the interpretative model promotes activating methods and involves a more inductive approach through which pupils gain insights that help them to construct theory. For this purpose, they must be given authentic and motivating tasks which are connected with familiar contexts. In this respect the interpretative model is consistent with the theories of Piaget and Vygotsky and – mutatis mutandis – with constructivist theory.

Pupil-pupil talk and learning by talk gains importance in the Low Countries

Influenced by the work of Barnes, Britton, and Rosen (1969), Barnes (1976), Barnes and Todd (1977) and by contributions at a number of national educational conferences in the early 70s, the 'Vereniging voor het Onderwijs in het Nederlands' and the 'Landelijke Werkgroep Moedertaalonderwijs en toch geen Nederlands' introduced ideas like 'learning by talk' and 'language in every school subject'. They did so by organising workhops, distributing small scale publications and by setting up school projects. Teachers as well as reseachers worked out topics like 'asking questions', 'oral interaction' and 'the integration of subjects'. In a very innovative six year school project called the 'Paulusproject', language didactics were gradually integrated in subjects like Biology and History, and later in Mathematics, Drawing and Economics. Teachers were introduced to dialogic teaching, collaborative learning, vocabulary learning strategies, etc and in doing so tried out team teaching *avant la lettre.* In the classroom, pupil-pupil talk became as important for learning as teacher-pupil talk. Decades ahead of its time, the project improved its educational quality through what would now be called an *integrated language policy* (van der Aalsvoort & van der Leeuw, 1982).

When we focus on collaborative learning, one of the first long term experiments with group work in the Dutch language took place even earlier, in the 1960s. The 'Werkgroep Empirische Introductie' (WEI) had a number of chemistry teachers organise group learning in a five year educational concept which emphasised on questions, assignments and texts. Pupils mainly processed the subject matter in groups and group discussions were the core of the teaching model. In the process pupils systematically increased their knowledge and understanding of the subject. One of the findings of the project was that this increase of cognitive understanding also affected the type of talk the pupils developed. As a matter of fact, language and learning appeared to go very much together (Leidse Werkgroep Moedertaaldidactiek, 1982). The Werkgroep Empirische Introductie also concluded that group work made it possible for teachers to pursue cross-curricular goals, such as social skills, learning-to-learn and citizenship. If pupils were to be given chances to learn by talking together, group work was an important means to do so (Leidse Werkgroep Moedertaaldidactiek, 1982). It may even improve the quality of interaction between teacher and pupils, as Brouwer confirmed more than twenty years later: 'Interactive, collaborative teaching methods have the benefit of improving the involvement of pupils in classroom activity and creating more horizontal interaction between pupils. In turn, by being more involved, pupils will be able to participate more actively in vertical interaction [with their teacher]' (Brouwer, 2006, p. 40).

Theory vs. practice in Flemish and Dutch classrooms

Realising the importance of interaction for learning is one thing, but implementing it in the classroom on a vast scale is another. Alarmed by international studies which found the systematic use of language for learning lacking in many classrooms, a number of Flemish and Dutch researchers decided to do a reality check. Geudens, Rymenans and Daems (1992) analysed the communicative patterns and use of language in four classes in secondary education (1^{st} and 4^{th} form general secondary education (Latin), transitional 1^{st} form B and 4^{th} form vocational education, carpentry) and for six different subjects (Biology, Geography, Technological Education, Social Education, Resource Study and Construction). They found that pupils' answers to teacher questions showed very little variation and followed a fixed pattern: teachers mainly asked closed questions. Questions which stimulated pupils to show their understanding of a subject matter or to formulate applications were scarce and questions urging them to make an analysis, to summarise or to formulate their personal opinion were hardly ever asked. As a result, pupils gave a lot of very short answers (one-word-answers mostly) and always a certain number of pupils remained silent throughout the whole lesson. Finally, teachers only gave their pupils feedback when their answer was wrong; they never gave any feedback when the answer was correct.

Geudens et al. (1992) confirmed international findings that pupils ask very few questions during lessons (see Chapter 2, section 2). The average in their study was one question every eight minutes, which comes down to a total of six or seven questions per lesson. Moreover, half of these questions had nothing to do with the subject matter. One third of their questions were requests for explanation of a given subject matter. Only in 17% of all cases did pupils ask for more information about a subject.

Interpreting these findings, the researchers explained this small number of pupils questions from the viewpoint of the teacher (he leaves little room for questions for fear of being delayed too much) and from that of the pupils (they consider asking questions a weakness - 'I am too stupid to understand this' - and/or they are focused on reproducing subject matter rather than on discussing it). Teaching methods also made a difference: pupils asked less for explanation when their teacher gave them clear instructions. They asked for more extra information when their teacher created an open, democratic climate.

Another finding of Geudens et al. (1992) was that pupils rarely made spontaneous remarks during their lessons, the average being one remark every ten minutes, or five per lesson. Most remarks - about 60% - had to do with the subject: checking if their answer to the question was correct, signalling difficulties while doing a task, explaining why they needed more explanation or pointing out links with their existing knowledge. A quarter of their remarks had nothing to do with the subject and mostly zoomed in on organisational aspects, such as homework, tests or things like 'What do we write in our diary, Miss?' A small percentage of their remarks was just irrelevant.

Summing up, Geudens et al. (1992) agree with the claim made by van der Aalsvoort and van der Leeuw (1992): unlike the use of language in daily life, language in the classroom is used in a very centralised way. *Centralised language use* means that the teacher does most of the talking while pupils are, in the first place, expected to just listen and pay attention, even as one of their fellow pupils is speaking. Pupils are only allowed to talk when their teacher permits them to do so. This rather limited discourse pattern quite typically illustrates a one dimensional view on education: language is nothing more than a vehicle needed to transfer knowledge, rather than a means to construct knowledge together. From this viewpoint pupils are prompted to reproduce knowledge instead of discussing or reconsidering it. This approach is very similar to Barnes' (1976) *transmission model* and is still common practice. Quite recently, for instance, Van Gorp (2010) found the same pattern in a kindergarten context and concluded that the quality of education in the classroom depends on the ways teacher and pupils talk to each other.

Teachers confront pupils with 'language barriers'

Analyses of pupil-teacher dialogue in primary education by van Gelderen (1994) and, fifteen years later, by Bossaert (2009), showed that teachers, without realising it, confront their pupils with 'language barriers'. Through his research, van Gelderen (1994) found that 8% of the words teachers used in their instructions were less frequent words. Also, the number of less frequent words increased grade by grade. Further, during lesson time, in 63% of the cases the teacher determined which pupil was allowed to answer his question. In 20% of the cases he directed his question to the class as a whole (or to no one in particular). In the remaining 17% it was unclear who was supposed to answer the question. Finally, of all questions asked, three kinds of answers were mainly required: 13% required the pupil to give a definition of some kind; 18% required that the pupil explained something and in 25% of the cases the pupil had to give a minimal answer (e.g. answer 'yes' or 'no'; complete an answer provided by the teacher; answer with one or two words; just nod). In only 12% of all question-answer exchanges did the teacher ask extra questions. Additionally, Bossaert (2009) found that teachers, without being aware of it, ask girls more factual questions whereas boys were more prompted to answer thinking questions. Also, their feedback on answers was of low quality (Bossaert, 2009).

Most recently, in the 2015 PISA report about collaborative problem solving, Flemish 15-year-olds showed good collaborative problem solving capacities on average, but when asked whether they had the opportunity to express their personal ideas during science lessons, only 37% said they could regularly do so. Wholeclass discussions of research projects (25%) and classroom debates about scientific issues (9,2%) occurred even less (OECD, 2016).

Research into practice of oral skills is too scarce in Flanders and the Netherlands

In 1997, in line with constructivist thinking, attainment targets and subsequent curricula in Flemish education confirmed and raised the status of communicative strategies, context bound learning and social and metacognitive skills (Ardui et al., 2011; Decoo, 2004; Vlaamse overheid, 2004). Teacher education followed suit and put strong emphasis on active learning and the implementation of (newly developed) cooperative strategies, thereby boosting the importance of speaking and listening skills. Surprisingly, however, results were poor. Systematic national assessments which measure the achievement of attainment targets in Flemish education (Vlaamse overheid, 2004, 2008, 2014) have pointed out that pupils are having much difficulty to acquire oral skills. In general, one pupil out of six in their final form of secondary education does not reach the attainment goals for listening and speaking. Specifically, half of the pupils find it difficult to summarise information, to deduce information from large amounts of information and to approach information with a critical mind. These last three deficiencies also showed in the national assessment studies for reading and writing (Vlaamse overheid, 2010).

So, what is the matter? As the attainment targets and curricula of multiple learning areas/subjects contain a variety of references to language skills, the problem must lie elsewhere. Research on this matter would be very helpful, but despite all stimulating projects and concept building mentioned above, studies on speaking and listening skills in education in Flanders and the Netherlands are and always have been scarce.

A clear illustration of this is the lack of recent relevant research in the field of Dutch as a school subject. In their reviews Hoogeveen and Bonset (1998), Bonset and Braaksma (2008) and Bonset and Hoogeveen (2011) found only a handful of studies on speaking and listening skills between 1969 and 1997, and as good as none between 1997 and 2008. Bonset and Braaksma (2008) believe this can be explained by a number of factors: first, compared with e.g. writing skills and from the teachers' point of view, teaching and assessing oral skills are very time consuming; second, oral skills are much more diverse than writing skills, which makes handling all aspects and finding the proper teaching methods and lesson materials not easy and very demanding; third, teachers seem to hold to the implicit idea that, as we are all born as listeners and speakers, teaching these skills is less essential or obligatory; finally, where learning goals for oral skills had a high social status in the 1970s, goals for writing skills have gradually taken their place in the 'goal hierarchy'.

To these explanations we wish to add the recent findings of Fisher (2011) who undertook a small-scale study among 75 British student teachers. Despite the promotion of dialogic teaching and speaking and listening skills in the British curriculum, practice seems reluctant to get along, Fisher observes. Previous research had already suggested that primary postgraduate student teachers bring some inflexible values and beliefs to initial teacher training programmes and a tendency to teach as they have been taught at school. Fisher found yet another mechanism that may cause (starting) teachers to keep dialogue out of their lessons, i.e. 'the tension between negative memories of classrooms characterised by 'all listening and no communicaton', which is the result legacy of low self-confidence in answering and raising questions. […] Students suffering from such cognitive dissonance are unlikely to develop sufficient confidence to create articulate classrooms' (Fisher, 2011, p. 33). In other words: teachers who were systematically kept silent as pupils, in turn risk keeping dialogue out of their classroom.

As a result of all these factors pupils do much listening (especially that) and speaking (or rather communicating) at school, in- and outside language classes, but little is being done to improve listening and speaking skills in their mother tongue. In many lessons, speaking and listening strategies are just absent. Knowing that the quality of speaking and listening can 'make or break' the quality of collaborative learning and that teachers do not seem to fully realise this (Kagan, 2014; Mercer, 1995), it is not surprising that they eventually skip activities like group work, as too often the only tangible outcomes are decibels.

Bonset and Braaksma (2008) conclude that only more research might put speaking and listening skills back on the educational agenda and raise their field status. As such not much seems to have changed since Lammers (1993) expressed his concern that there is no real tradition in teaching oral skills in Flemish and Dutch education, nor is their any coherent framework of knowledge and insights concerning didactic methods and assessment strategies. This, Lammers concluded, drops us into a vicious circle, for without a theoretical and didactic base teachers will not feel inclined to pay much attention to oral skills, while in turn leaving researchers on bare ground to conduct e.g. effect studies of current practices. It is not surprising, Bonset and Braaksma (2008) add, that the lack of instruments to assess oral skills has brought instrumental research to a standstill.

Contrastively, there is abundant international research on speaking and listening skills, which focuses on either teacher-pupil or pupil-pupil talk, or both. This will be discussed in Chapters 2 and 3.

Exploratory talk in the Flemish and Dutch educational context

This dissertation zooms in on a type of classroom talk which has not yet been highlighted in Flemish and Dutch research until very recently and which is also unknown to the vast majority of the teachers. More importantly, we believe it may be a missing link in the listening and speaking didactics in Flanders and the Netherlands and help teachers to make dialogic teaching and especially collaborative activities more worthwhile. It is called *exploratory talk* and it was first reported by Dawes, Fisher and Mercer (1992). Shortly afterwards the concept was also introduced in the Dutch-speaking countries, though briefly, by van der Aalsvoort and van der Leeuw (1992):

> Een mogelijkheid om dat eenrichtingsverkeer in het taalgebruik in de klas te doorbreken is het werken in groepen. Daarin kunnen voorwaarden voor leerlingen geschapen worden waardoor ze meer en ontspannen kunnen praten en niet zozeer leren door luisteren, maar leren door praten. Gecentraliseerd taalgebruik is in groepswerk vervangen door aarzelend en exploratief taalgebruik. In een sfeer van tolerantie, van zoekende formuleringen en van gemeenschappelijke pogingen kunnen leerlingen hun taalgebruik uitbreiden door hun eigen gedachten en de meningen van anderen met elkaar te combineren. (van der Aalsvoort & van der Leeuw, 1992, p. 45)
>
> ['An opportunity to break the one-way traffic in the language of the class is working in groups. Group work creates conditions for pupils to talk more and in a relaxed way rather than having to learn by listening. In group work centralised language is replaced by hesitant exploratory language. In an atmosphere of tolerance, of seeking formulations and common efforts, pupils can expand their language by combining their own thoughts and opinions with those of others.']

Depending on the teaching method, discussions in the classroom are organised in two possible ways, van der Aalsvoort and van der Leeuw explain, i.e. teacher-pupil or pupil(s)-pupil(s). These discussions should have an *exploratory nature*: pupils find and explore a subject or topic using language until they find answers to questions, solutions to problems, arguments for opinions. They do that by talking and writing. They listen to and read texts in an exploring way. They connect new knowledge with their own existing knowledge and experience. Only after the exploratory phase do they check out if what they think is correct. The way they use language here is very similar to the way language plays a role in everyday life and learning (van der Aalsvoort & van der Leeuw, 1992).

With this explanation van der Aalsvoort and van der Leeuw refer to characteristics of exploratory talk described by Barnes (1976), who initiated the concept, and to Mercer (1995) who defined it and conducted further empirical research on exploratory pupil-pupil talk during collaborative activities in the classroom.

Related research topics in Flemish and Dutch research

Whereas exploratory talk has become the subject of much international research, it has been pretty much ignored by Flemish and Dutch researchers, a few exceptions notwithstanding. We did a brief search of the Dutch equivalents of the concept ('Exploratief gesprek', 'Explorative gesprekken', 'Explorerend gesprek', 'Explorerende gesprekken') in Google Scholar in order to find more about exploratory talk in other Dutch sources, but did not find any further references within an educational context. Nevertheless, we found some interesting Flemish and Dutch research concerning dialogic teaching and collaborative learning. We will briefly discuss five illustrative studies.

Van den Branden and Van Gorp (2000) introduced the CLIM method (Cooperative Learning in Multicultural Groups, see Paelman, 2004; see also Kagan, 2014) to small groups of 11- and 12-year-old pupils in two primary schools. The CLIM method - or tool - answers to the need of teachers for strategies that would enhance language skills of non-Dutch speaking pupils. The tool focuses on participative equity, mutual responsibility and group ownership, but also on constructive content interaction. Shortly after its introduction it was found to be beneficial to all learners in a classroom context. The intervention by Van den Branden and Van Gorp (2000) did not comprise an introduction to any strategy for group talk, but qualitative analysis of talk demonstrated that a task with a clear goal, like a well-defined problem, not too close and not too open, allowed the pupils the necessary intellectual and creative freedom to 'talk themselves into the solution'. In short, it improved their problem solving capacities through talk.

Damhuis and De Blauw (2006) developed an instrument for language development through interaction. One of the means to develop language skills, they considered, is active learning through talk. In an interventional study they had teachers combine active learning with effective talk, thus combining cognitive goals with language goals. They found that for talk to be effective, i.e. to have learning potential, pupils must be given opportunities to talk in pairs or in small groups, the topic of discussion must be relevant and challenging, and they should receive feedback on their interactive processes. Teachers may hesitate to give pupils such opportunities, but 'Even as they are young, children can be challenged to do complex thinking and talking: compare, reason, argue, conclude' (Damhuis & De Blauw, 2006, p. 16).

Hoek (2007) had two mathematics teachers train their 16- to 18-year-old students in strategies for interactional learning and analysed their talk patterns within small groups for a year. He found that the students developed a more critical attitude towards their peers, enriched their conversation with more new input and explored more possible solutions to the mathematical problems they had had to tackle. Their type of talk evolved from rather cumulative to exploratory.

Though the focus of his study was Dutch as second language learning, Van Gorp (2010) found conclusive evidence for the importance of dialogic teaching and for active, collaborative learning as a means to co-construct knowledge. The researcher found that the quality of interaction between teachers and their pupils on the one hand, and among pupils on the other, was low. Moreover, it was very much determined by the intrinsic qualities of the teacher and by the extent to which he was aware of the importance of effective interactions. Van Gorp (2010) concluded that, when it comes to constructing knowledge, pupils often lack scaffolding.

Finally, van der Veen, Dobber and van Oers (2017) describe positive outcomes after a large scale intervention in 11 Dutch kindergarten schools, to promote young children's oral communicative competence through 'productive classroom talk'. They did so by using a model called MODEL2TALK. This model promotes the use of several teacher talk moves (or tools) that can be used to encourage children to share their ideas, listen to one another, reason, think together and reflect on their communicative performance.

These five studies illustrate genuine academic interest in the role of language for learning in both Flanders and the Netherlands, especially in forms of dialogic teaching. At the same time, a thorough word search showed that the concept of exploratory talk itself – or a synonym - is mentioned explicitly in neither the Flemish attainment targets nor in the various curricula (T'Sas, 2013). Characteristics of exploratory talk can be found in these sources but, as they are distributed over several subjects, like Languages, Social Skills and Man and Society (part of the former World Orientation), they lack coherence. In the curricula we examined we found no mentioning of a didactic learning trajectory for exploratory talk or similar (GO!, 2013a; Vlaams Verbond van het Katholiek Onderwijs, 2013). Also, exploratory talk is not mentioned in the annual 'Onderwijsspiegel', which summarises findings of the official school inspectorate (Vlaamse-Ministerie-van-Onderwijs-en-Vorming, 2015; Vlaamse overheid, 2011). This suggests that absence in attainment targets and curricula also means absence in classroom practice.

2. Research questions

The previous section showed that speaking and listening skills in the Flemish and Dutch educational context deserve more academic and practical interest. Exploratory talk as a means to improve learning and to raise the quality of group work seems as good as uncharted territory in the Low Countries. This is problematic, because at present, there is no way of knowing whether exploratory talk can be implemented in Flemish classrooms as it has been in other countries, whether this implementation would add to the educational value of collaborative learning or which learning outcomes it might generate. Though we strongly believe this in itself is a major reason to 'fill the research gap', we would like to substantiate the need for the present study even more.

As will be described in the next chapter, international research on exploratory talk has found and confirmed its educational potential in collaborative activities: pupils improve their reasoning skills, work better together and also get better at solving problems, both at group and individual level. Also, teachers discover the added value of dialogic teaching by modelling exploratory talk themselves in order to teach it to their pupils and by observing the results in their practice. If findings like these can be confirmed in the Flemish educational context, they may offer teachers extra resources and boost speaking and listening pedagogy.

Furthermore, based on findings in international research exploratory talk can also be an important tool in the language policy of a school (Mercer, 1995; Rojas-Drummond, Perez, Velez, Gomez & Mendoza, 2003; Setati, Adler, Reed & Bapoo, 2002; Wegerif, Linares & Rojas-Drummond, 2005). This is very important to know, as every Flemish school must realise a language policy (cf. infra), which Van den Branden (2012) defines as following:

> ‘Taalbeleid is de structurele en strategische poging van een schoolteam om de onderwijspraktijk aan te passen aan de taalleerbehoeften van de leerlingen met het oog op het bevorderen van hun algehele ontwikkeling en het verbeteren van hun onderwijsresultaten’ (Van den Branden, 2012, p. 11)
>
> [Language policy is the structural and strategic attempt of a school team to adapt its educational practices to the language and learning needs of its pupils, in order to improve their general development and their learning outcomes.]

This definition implies that schools focus on language pedagogy and communicative strategies, in and beyond language classes. To that end both the Flemish and Dutch government have taken several initiatives. In 2006 the Nederlandse Taalunie developed a referential framework which sums up the language competencies every teacher in the Dutch speaking countries should master (Paus, Rymenans & Van Gorp, 2006). The framework describes these competencies from within thirteen different sets of goals, each of which relates to the professional roles teachers play. Examples of these goals are: talking with learners, formulating oral and written instructions, making text materials accessible to learners, etc. The framework contains many concrete examples how kindergarten, primary and secondary teachers can meet these goals. And in 2007 the Flemish government imposed on all Flemish schools the obligation to implement a language policy (Vandenbroucke, 2007), while the quality of this implementation would be analysed regularly by the Inspectorate. Since then more initiatives have been taken to encourage schools, offering them tools to develop a tailor-made language policy, but the quality of this policy differed (and differs) substantially from school to school (Vlaamse overheid, 2011). This may be a hindrance to the implementation of exploratory talk for learning. We are convinced it can contribute to the quality of any language policy, as it can be embedded in almost any subject, but a lot depends on the school’s policy-making. Nevertheless, this study may stimulate schools to actually implement exploratory talk.

Another reason for this study is the introduction of research competencies in the specific Flemish attainment targets in 2007. Part of these competencies are skills which come very close to the characteristics of exploratory talk which will be described more exhaustively in Chapter 3: asking the right questions, synthesising information and drawing conclusions. Research competencies have gradually been introduced in the curricula. In a first stage, pupils acquire skills and attitudes that allow them to work more independently and self-regulatory: ‘In the first grade[4] research competencies offer pupils many chances to find information in an efficient way and to produce arguments to support this’ (GO!, 2013a, p. 3). Group work is mentioned as a strong learning tool for integrating these Competencies into every subject and

---

[4] In Flemish education the word ‘level’ is used to refer to a period of two consecutive years (‘leerjaren’ or forms) in primary or secondary education. Primary and secondary education each comprise six years or three levels. Pupils start in level 1 and gradually move up to levels 2 and 3.

to us it is obvious that exploratory talk can prove its added value in this. Therefore, this study may offer teachers inspiration to help develop their pupils' research competencies.

Finally, as exploratory talk is a 'great unknown' in Flemish and Dutch research, the same holds for its implementation in the teacher training curriculum. If the results of this study can confirm its added value for education, as has been demonstrated in international research, there will be scientific ground for embedding it in teacher training practice and – mutatis mutandis – in general classroom practice.

Summarising, the purpose of this study is to empirically test the implementation and potential for learning of exploratory talk in group work, answering the following research questions:

> RQ 1. What is exploratory talk, how is it measured and which effects does it have?
>
> RQ 2. To what extent do pupils of the third level (primary school) use exploratory talk in group assignments?
>
> RQ 3. To what extent do pupils use exploratory talk after a 12 week training?
>
> RQ 4. What effects does the use of exploratory talk have on pupils' problem solving skills at group and at an individual level?
>
> RQ 5. Which variables explain evolutions in the individual use of key words for exploratory talk?

3. Research design

In order to answer these research questions, we carried out two studies: a narrative review into the definition, the measurement and the effects of exploratory talk (Study 1) and a quasi-experiment with pre- and post-testing (Study 2). The following figure shows the global research design as well as the link between the studies and the research questions.

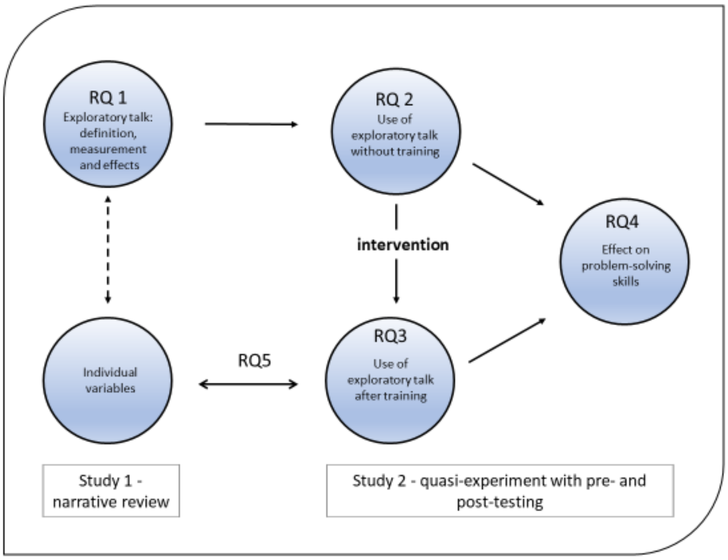


Figure 1. Global Research Design Model

In this diagram we put each research question in a sphere (except for RQ 5, which lies in between two spheres). Different kinds of arrows express the relationship(s) between these spheres. In order to answer these questions two studies took place: a narrative review (Study 1) and a quasi-experiment (Study 2, cf. the rectangles in the Design Model). The black arrows express a research timeline. The dotted arrow expresses interaction between research activities: the narrative review is to provide information about individual variables worth examining, starting with RQ 1.

The timeline goes like this:

Study 1: by answering RQ 1 we want to clearly determine the definition, measurement and effects of exploratory talk. We also want to identify potentially influencing individual variables. A narrative review must provide this information.

Study 2: the outcome of the narrative review is necessary in two ways: a) to proceed to the next research step, i.e. to construct and carry out the quasi-experiment which enables us to answer RQ 2 and after that RQ 3 and RQ 4; and b) to check the importance of certain individual variables that need to be investigated during the experiment (intervention) and which are necessary to answer RQ 5.

The answers to RQ 2 reflect the outcomes of a pre-test and therefore have to be obtained before the intervention. The intervention in turn leads us to the post-test, based on which we can answer RQ 3. With the information obtained after answering RQ 2 and RQ 3 we can answer RQ 4. The information obtained by answering RQ 5 helps to complete the answers of RQ 3.

Chapter 2: Theoretical framework

The subject of this thesis is the importance of language for learning and, more specifically, the importance of a specific type of talk, i.e. *exploratory talk for learning* in the classroom, a concept which was launched by Douglas Barnes as early as 1976. Therefore, we believe it is necessary to walk down the path from which this idea has come and integrate the context for this study into a theoretical framework which is based on a diversified literature study, before giving a literature review on exploratory talk itself. In this chapter we first look at the role of language in modern learning theories (1). Second, we zoom in on the use of language for learning in the classroom (2). Third, further narrowing down the scope of this topic, the use of language for learning in collaborative activities, i.e. group work is examined (3).

1. Learning theories and the role of language

How important is language for learning? In classical antiquity Socrates asked his students questions that made them aware of the weaknesses in their thinking. Handling their master's feedback, they learned by their mistakes. Today, in constructivist thinking, the Socratic dialogue is again an important tool in learning and evaluating practices (Lipman, 1981; Topping & Trickey, 2014). But between then and now the importance of language for learning had to be rediscovered, starting at the beginning of the 20th century.

The behaviourist view

In the early 20th century, behaviourists like Watson and Skinner translated learning - including the learning of language - into a system of observable behaviour, induced by proper stimuli (Valcke, 2010). Internal processes that take place in the mind were not verifiable and therefore not taken into account. Moreover, mental processes, like knowing, thinking, believing, etc, were considered as forms of latent, non-observable behaviour which respond to external or internal stimuli (Caroll, 1970). Learning a language, too, was a matter of improving one's responses to repeated stimuli. This view fortified a transmissional view on education: a teacher's main task was to transmit responses that were appropriate to a certain stimulus and to transmit stimuli in order to evoke proper responses. This would encourage learners to keep adjusting their behaviour until they received positive reinforcement (Skinner, 1972). It is no surprise, then, that from this viewpoint, programmed step-by-step instruction, 'skill and drill' exercises, question and answer frameworks of gradually increasing difficulty and so on were considered to be the main ingredients of a teacher's repertoire while language was merely one of the possible ways to transmit information (stimuli).

The cognitivist view

Cognitive psychologists, like Bruner and Ausubel, did not agree with the rather reductionist thinking of behaviourists (Valcke, 2010). They wanted to know what went on inside the 'black box', i.e. which mental processes took place during learning. For cognitive psychologists, learning is processing information (Van Petegem, Imbrecht & Brandt, 2005) and language is one of the many human mental or cognitive activities. Primarily, cognitivists studied how learners process information and how this information is stored away in memory. Learning became the result of mental constructions made by the learner. Later research was extended to topics like memorising, learning, thinking and problem solving. Subsequently, researchers went on to investigate what role existing knowledge plays in the acquisition of new knowledge (Van Petegem et al., 2005). Knowledge was soon thought to be actively constructed by the learner rather than passively absorbed. According to Bruner, for instance, learners select information, code it into mental representations and process these in order to integrate them in their existing cognitive structure (Bruner, 1962). By consequence, teachers should connect new subject matter to their pupils'/students' current knowledge as much as possible. In order to increase insight and memory storage they should also reintroduce subject matter in a progressively more complex way ('spiral curriculum'; Bruner, 1960). Bruner further postulated that the process of knowledge acquisition requires active *discovery learning* with a strong emphasis on problem solving skills. Only then can education fully stimulate a child's cognitive development.

David Ausubel introduced the concept of *meaningful learning*, i.e. the process of connecting new subject matter with known information and contexts (Ausubel, 1960). This looks similar to Bruner's ideas, but Ausubel did not share Bruner's preference for discovery learning. Instead he preferred *receptive learning*: information is presented to learners in such a way that they have little difficulty integrating it into their cognitive structure. This does not mean that learning becomes a passive process, it remains active. Teachers must find ways to encourage active involvement with the entire learning process. They can facilitate meaningful learning by using *advance organisers*, which represent the structural hierarchy of knowledge. Advance organisers are relevant subsuming concepts that help learners to organise new subject matter. They can be very general and abstract or very specific and concrete (Ausubel, 1960; see also Bullen, Moore & Trollope, 2002).

Thus, a teacher's role is not to drill knowledge into learners through consistent repetition (except perhaps for the memorisation of facts and formulae), or to work with rewards and punishments. Rather - and here the ideas of Bruner and Ausubel are joined - the teacher should facilitate discovery by providing the necessary resources and by guiding learners as they attempt to assimilate new knowledge to existing knowledge and to modify the old to accommodate the new. This way, compared to behaviourism, language received an extra function, i.e. that of a framework for constructing and integrating knowledge. Later acknowledgment of this view came from e.g. James Britton. In his classic study *Language and Learning* we read: 'Language is our principal means of classifying, and it is this classifying function that goes furthest towards accounting for the role of language as an organiser of our representations of experience' (Britton, 1970, p. 8) and 'We cannot underestimate the value of language as a means of organising and consolidating our accumulated experience […]' (Britton, 1970, p. 319).

The constructivist view

In line with Bruner, Jean Piaget and Lev Vygotsky developed a constructivist view on learning (Perry, 1999; Valcke, 2010). Constructivists consider knowledge not as something that is acquired, but as something that is actively constructed by learners. In doing so, learners use their existing cognitive structures, i.e. their personal experiences and hypotheses of the environment, which they continuously test through social negotiation. They execute cognitive functions such as memorising, structuring, analysing, etc (Vermunt, 1992). Also, they regulate their own learning process through metacognitive activies such as monitoring, assessing and concentrating (Vermunt, 1992). This means that learning is an individual and intentional process, but it is also a social process and this is where language comes in. De Corte (1996) summarises all constructivist characteristics of learning in the following definition:

> 'Leren is een constructief, cumulatief, zelfgestuurd, doelgericht, gesitueerd, coöperatief en individueel verschillend proces van kennisverwerving, betekenisgeving en vaardigheidsontwikkeling' (De Corte, 1996, p. 145)
>
> [Learning is a constructive, cumulative, self-regulatory, intentional, context bound, collaborative and individual process of acquiring knowledge, giving meaning and developing skills.]

Cognitive constructivism

The word 'individual' in De Corte's (1996) definition implies that learning is relative to the learner's stage of cognitive development. This makes understanding his existing intellectual framework central to understanding his learning process. For that reason, Piaget developed theories of childhood development and education. As said before, Piaget did not believe that learning was the passive assimilation of given knowledge. To him, learning is a dynamic process in which we actively construct knowledge by creating and testing our own theories of the world (Piaget, 1970). And because this process is also something we must acquire while growing up, it comprises successive stages of adaptation to reality.

A view on Piaget's theories which we believe to be interesting for our study was added by William Perry who acknowledged that learners adapt and develop by assimilating and accommodating new information into existing cognitive structures (Perry, 1999). He also believed in the 'Piagetian hierarchy', the idea that one cognitive structure builds upon the previous one. To Perry, though, Piaget's theories lacked consideration about the context of learning: learners do not approach knowledge from the same point of view. They are all influenced by their gender, race, culture and socioeconomic background. Moreover, learners have the flexibility to look at the world from varying points of view. Therefore, the developmental process is a constantly changing series of transitions between various positions (Perry, 1999). Perry claimed that learners 'journey' through nine 'positions' with respect to intellectual (and moral) development. These stages can be characterised in terms of the student's attitude towards knowledge and this attitude is expressed by language (Perry, 1999). This way, Perry seems to go along with Dewey's focus on metacognition which would allow learners to build knowledge onto prior experiences and preconceptions (Dewey, 1933).

In Piagetian terms, we all learn how to assimilate and accommodate new information. The constructivist teacher, then, must employ a number of strategies which help learners to actively acquire knowledge. At the same time, the teacher must stimulate the learner's metacognitive skills. If learners are to become aware of the boundaries of their existing knowledge and to accept the need to modify or abandon existing beliefs, they cannot do so without being able to reflect on their personal skills and knowledge. Very important, cf. De Corte's (1996) definition, is that the learner has the intention to do precisely that and unlike what the behaviourists claim, external rewards and punishments such as low grades or extra homework will not be sufficient (Perry, 1999).

Social constructivism

Whereas Piaget may be considered one of the pioneers of individual constructivism, post-revolutionary Soviet psychologist Vygotksky may well be called the first social constructivist (Valcke, 2010). Social constructivism emphasises the collaborative nature of (much) learning.

Vygotsky rejected the assumption made by Piaget that one could separate learning from its social context. His view approaches Dewey's conviction that education and learning are social and interactive processes (Dewey, 1933). He argued that new cognitive development originates in, and must therefore be explained as products of social interactions. In other words, whereas Piaget thought new concepts of reality arose from solitary reflection after engagement with others (Bullen et al., 2002), Vygotsky believed we learn through collaborative interaction. Hence, learning cannot simply be the assimilation and accommodation of new knowledge by learners; it is the process by which learners are integrated into a knowledge community. Consequently, Vygotsky (1978) claims, learning first takes place socially, between people (intermental) and later individually (intramental). All the higher functions in our brain originate the same way: as actual relationships between individuals and primarily, this process is mediated by language.

Vygotsky (1978) also made a distinction between two developmental levels: the level of *potential* or *proximal development* is the level at which learning takes place. It comprises cognitive structures that are still in the process of maturing, but which can only mature under the guidance of or in collaboration with others. The level of *actual development* is the level of development that the learner has already reached. It is the level at which the learner is capable of solving problems independently. For teachers, Vygotsky's theories mean they should focus on and give structure to collaborative learning methods which require learners to develop teamwork skills and to see individual learning as essentially related to the success of group learning.

The educational neuroscientific and neuropedagogical view

Over the last decades there have been discussions about the values and research base of constructivist thinking in education (for a review, see Green & Gredler, 2002). At the same time, characteristics of constructivist learning have become the research subject of neurosciences like neurobiology, neurophysiology and neuropsychology (Van Camp, Vloeberghs & Tijtgat, 2015; Willis, 2007). Neuroscientific researchers use modern technology, like MRI and fMRI-scans, EEG's, etc, to find out which neurological processes occur during learning and how education can benefit from those insights. Cognitive neuroscientists study the various networks in our brain, more specifically the ones that show activity as we perform certain tasks. They study the way in which those networks change as we receive input, e.g. while we are reading, listening to a lecture or watching a movie. This has led to the development of a new scientific branche called *educational neuroscience* or *neuropedagogy*. Educational neuroscience is interdisciplinary, i.e. it brings together researchers from (cognitive) neuroscience, instructional psychology, pedagogy and other related disciplines in order to investigate the interaction between biological processes and education (Howard-Jones, 2010, in Van Camp et al., 2015). The outcomes of their work have created new insights in learning processes which may be beneficial for both teachers and learners. For instance, a connection was found between the brain systems that implement deductive reasoning, memory and language tasks (Barbey & Barsalou, 2009).

One of the main learning principles in educational neuroscience is active learning, if only because it 'takes advantage of processes which stimulate multiple neural connections in the brain and promote memory' (Kaufer, 2011; see also Willis, 2007) and because it can raise the quantity of spoken language in the classroom (Mercer, 1995). In order to understand this, we can use Bloom's taxonomy of lower-level and higher-level cognitive functions. Lower-level functions are e.g. understanding and remembering, while creating, evaluating, analysing and applying are higher-level functions (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956). The latter are the more complex thought processes. Neuroscience has revealed that each of these sets of functions is associated with different parts of the brain. Lower-level functions are associated with the area of the brain (the hippocampus) that regulates our memory and special awareness. Higher-level functions are associated with the areas of the brain (the cortical areas) that motivate us and that regulate our ability to make decisions and associations. As higher-level functions involve a greater number of neural connections and more neurological 'cross-talk', they are more beneficial for learning. Active learning, then, stimulates a variety of areas in the brain and promotes memory (Van Camp et al., 2015).

Additionally, Tokuhama-Espinosa (2010) has made clear that using language involves higher-level or higher-order thinking. Connecting both ideas, Van Camp et al.'s (2015) and Tokuhama-

Espinosa's (2010), brings us very close to understanding the learning potential of higher-order talk (Sutherland, 2006). Sutherland refers to the study on reciprocal reading by Palincsar and Brown (1986) and to the experiment with guided reading of Pressley and Allington (2014). Both have demonstrated how higher-order talk can be generated by group work in which teacher initially models and enables pupils to practice skills of questioning, predicting and inferring. Doing so, he gradually moves them towards independence. Exploratory talk, as Sutherland (2013) found in her own study, contains essential indicators of higher-order thinking, which makes it a relevant tool for education (see Chapter 3). This was confirmed by Soter et al. (2008) who did a study on talk and indicators of high-level comprehension. Their data indicate that 'the most productive discussions are structured, focused [and] occur when students hold the floor for extended periods of time […]. Results also indicate that authentic questions give rise to longer incidences of student talk, which in most cases result in opportunities for greater elaboration of utterances by students, and which in turn, generate reasoning and high-level thinking' (Soter et al., 2008, p. 373).

Two predominant views on the role of language in learning

The overview of modern learning theories we have presented so far shows that it has become very difficult to imagine learning without language. Hence, it is not surprising that the correlation between language and learning in the classroom has been the subject of multiple research. It has inspired researchers to reconceptualise and rename it as e.g. *interthinking* or the process by which people use language to think creatively and productively together (Littleton & Mercer, 2013). Considering this, there has been a long discussion among researchers about which theory to lean upon when studying the role of language in learning (Bullen et al., 2002). Two views stand out: the (neo-)Piagetian view and the (neo-)Vygotskyan view. We do not want to go into this matter in detail, but in brief, this is what this discussion is about:

The Piagetian view

For Piaget, language is 'a system of symbols for representing the world as distinct from actions and operations that form the processes of reasoning' (Wood, 1998, p. 25, cited in Mercer & Littleton, 2007). Piaget agreed with cognitive psychologists that learners' active construction of their own understanding was fundamental to their cognitive development. He rejected the transmissional model for learning (Barnes, 1976) in which knowledge is passed from adult to child. Moreover, as he considered cognitive development to be an individual process, he considered interaction with adults as a possible hinderance to this development. Instead he found it very important that adults, including teachers, provide a rich, stimulating and generally supportive environment for children. On the other hand, Piaget did find interaction between children a potential source for learning, thereby opening the door for didactic approaches such as group work and peer collaboration (Atwood, Turnbull & Carpendale, 2010). As social interaction can take many forms, one being more fruitful for learning than the other, Piaget argued that those forms of social interaction that facilitate development must be specified. He distinguishes between relationships/interactions of *cooperation* and *constraint*, where cooperation leads to constructive talk and understanding, and constraint inhibits those, even though in reality both relationships/interactions are mere extremes on a continuum.

The Vygotskyan view

Vygotsky was one of the first to emphasise the role of language and culture in cognitive development. According to Vygotsky, language and culture play essential roles both in human intellectual development and in how humans perceive the world. They are the frameworks through which humans experience, communicate, and understand reality. Language, then, is both a psychological and a cultural tool. As a psychological tool it helps learners to make individual sense of experience. As a cultural tool, it is used to share experiences and make sense of them in a collective way. As a result, human cognitive structures are, so Vygotsky believed, essentially socially constructed. In short, knowledge is not simply constructed, it is co-constructed (Vygotsky, 1978) and 'the development of children's thinking is shaped by the dynamic relationship between intermental activity (social interaction) and intramental activity (individual thinking), with language as the prime mediator between the two' (Mercer & Hodgkinson, 2008, p. 10). Vygotsky's work has laid the foundation for the so-called sociocultural theory which explains cognitive development and learning in a cultural and social context. It focuses especially on the dialogic approach of learning. 'It studies how people use language as a social mode of thinking' (Mercer, 1995, p. 4).

A possible reconciliation of both views: Chapman's epistemic triangle

Piaget was long criticised for not taking into account the role of social interaction in development but that criticism may be unfair. Quoting Bruner, Atwood et al. (2010) refer to a 'commonly held assumption that Piaget ignored the role of social factors in development' (Atwood et al., 2010, p. 360). Hence, some researchers considered Piagetean theory of little use when it comes to studying the role of teaching and classroom talk in learning. Atwood et al. (2010) refer to Chapman (1991) who integrated the Piagetian and the Vygotskyan theory in his concept of the 'epistemic triangle' to reconcile both views. The basic idea behind this triangle is that knowledge develops in and through one's interactions with the world on the one hand, and with one's social interactions with others about each's experiences of the world on the other. In education 'the world' is a collection of cognitive and metacognitive activities in the classroom, while social interaction takes place between the teacher and his pupils or between pupils. This suggests that there is interaction between the interpersonal and the intrapersonal, which comes close to Vygotsky's theory. Atwood et al. (2010) conclude: 'Much of the current research on the role of social interaction and perspective taking in classroom learning is consistent with the epistemic triangle and with a Piagetian view of the child's active role in the construction of knowledge' (Atwood et al., 2010, p. 361).

We did not find evidence that Chapman's reconciliation of Piagetian and Vygotskyan views has been widely acknowledged, but perhaps it all comes down to perspective. Mercer (1995) sees a functional and perhaps even pragmatic distinction between Piaget's and Vygotsky's theories, referring to the former as the study of the origin and nature of knowledge, and to the latter as the study of the process of teaching and learning (Mercer, 1995). Though Piaget may well leave room for social interaction, 'Vygotsky's theory, more than Piaget's, has room in it for teachers as well as learners. It draws the attention to the construction of knowledge as a joint achievement [and provides] us with a theory of the development of thought and language' (Mercer, 1995, p. 72-73).

Based on international research data Alexander (2008), too, separates Piaget's thinking from Vygotsky's by allocating different pedagogical values to them: teaching as facilitation (Piaget) and teaching as acceleration (Vygotsky). Teaching as facilitation implies the application of developmental principles: 'The teacher respects and nurtures individual differences and waits until the children are ready to move on instead' (Alexander, 2008, p. 98). Teaching as acceleration means that 'education is planned and guided acculturation [in which] the teacher seeks to outpace development rather than follow it' (Alexander, 2008, p. 98). Emphasising the important role of the teacher, which seems less important in Piaget's theories, and the dynamic, dialogic nature of the learning process, Mercer has proposed a new neo-Vygotskyan concept in which he joins the concepts of scaffolding and the zone of proximal development (Mercer, 2000; Mercer & Hodgkinson, 2008; see also Mercer & Littleton, 2007; Littleton & Mercer, 2010). He calls it the *Intermental Development Zone*. It is a zone, or a bubble, created by language in which either teacher and pupil(s) or two or more pupils reason about and develop common knowledge. It 'represents a continuing state of shared consciousness, focused on the task in hand and dedicated to the objective of learning' (Littleton & Mercer, 2010, p. 110-111).

Building largely on Vygotsky's theories, many sociocultural researchers and educators have promoted the collaborative use of language in the classroom (Barnes et al., 1969; Britton, 1970; Barnes, 1976; Howe, 1992; Bowskill, 2010; Coultas, 2012; Enghag, Gustafsson & Jonsson, 2009; Harris & Ratcliffe, 2005; Kerawalla, Petrou & Scanlon, 2013; Mercer, 1995; Rojas-Drummond, Torreblanca, Pedraza, Velez & Guzman, 2013; Sutherland, 2006; Alexander, 2004). Together with Carter (2002) many agree that talk is undervalued in education, leaving the classroom scene to skills or subjects which have a higher status, such as writing and literature.

Research on the collaborative use of language in the classroom focuses on either teacher-pupil talk or pupil-pupil talk, and sometimes on the correlation between both (for an overview see Chapter 3, section 3.1.2). From a Vygotskyan point of view both foci are equally important.

In teacher-pupil talk the teacher plays an important guiding role in the learning process. First, the teacher is the adult, without whom pupils may not be able to master knowledge within the zone of proximal development. His role is to scaffold their learning process. Second, as a scaffolder, the teacher also functions as a role model for his pupils when it comes to using that type of language which has more learning potential, such as dialogic teaching (Alexander, 2004; Mercer & Hodgkinson, 2008). Although a number of researchers found that the use of more dialogic strategies achieves better outcomes for children (Mortimer & Scott, 2003; Rojas-Drummond & Mercer 2003; Alexander, 2004; Mercer & Sams, 2006; Jay et al., 2017), practice shows that there are very different ways for teachers to use classroom dialogue when interacting with children. Variations and the extent to which dialogic teaching takes place appear to be teacher but also culture or country dependent (Alexander, 2004).

Investigating pupil-pupil talk is equally important, because the context of learning differs fundamentally from the context of teacher-pupil talk. On the whole, literature shows considerable support for the idea that knowledge is a co-constructed activity of pupils and that collaborative talk can support them to develop higher-order thinking, high-level understanding, the voicing of personal opinions and ideas, and argumentation skills (e.g. Fernandez, Wegerif, Mercer & Rojas-Drummond, 2001; Harris, 1995; Lofgren, Schoultz, Hultman & Bjorklund, 2013; Mercer, 1995, 2013; Mercer & Littleton, 2007; Mercer & Sams, 2006; Mercer, Wegerif & Dawes, 1999; Rojas-Drummond & Mercer, 2003; for an overview, see also Sutherland, 2013).

As our research focuses on pupil-pupil talk, we will go further into this matter in section 3 of this chapter, but first we summarise what is known about the factual use of language for learning in the classroom.

2. The use of language for learning in the classroom

In his theories on the relationship between language and learning (or language and education), Halliday (1979) distinguishes between learning language, learning through language, and learning about language. For the purpose of this dissertation we will only consider learning through language and look at how this is stimulated - or inhibited - by classroom practice.

The process of learning through language

According to Halliday learning in classrooms, irrespective of the subject which is taught, is primarily accomplished through language. Teachers pass on information, give demonstrations, explain subject matter, ask open and/or closed questions, organise or lead discussions, assign reading and writing tasks. Pupils read texts, take notes and write essays, listen to what the teacher tells them, take in information provided by media or listen to presentations by their peers, answer teacher questions, participate in whole-class and peer group discussions, etc (Halliday, 1979). As we saw previously, there is general agreement that social interaction must play a role in the development of knowledge, but which forms of classroom interaction/language practice facilitate which kinds of learning? Halliday (1979) found three language practices of interest to educational researchers, i.e. scaffolding, sharing time and the I-R-F sequence. Though this dissertation focuses on pupil-pupil talk we find it important to highlight the role of the teachers first, as they play a key role in facilitating and supporting such talk.

Scaffolding, sharing time and the Initiation-Response-Feedback (I-R-F) sequence

Scaffolding is the process through which an adult supports a child by helping it to perform a task that they would not be able to master on their own. He or she can do so by explaining strategies, evoking reflection on previous work, ask elaborative questions, etc. Or, as Atwood et al. (2010) state: 'Optimal conditions for verbal reasoning are those in which student talk is facilitated and carefully managed by teachers who not only give students ample time on the floor but who manage their own questions and comments well' (Atwood et al., 2010, p. 398).

The term 'scaffolding' was introduced by Wood, Bruner and Ross (1976), who made clear that pupils/students can act as the 'adult' learner, too. As a matter of fact, they can be experts at different points in a task (Tin, 2003). Exploratory talk, the subject of this study, relies very much on forms of scaffolding. This will further be discussed in the next chapter.

Sharing time, also known as show-and-tell, is one of the few activities in which children are expected to talk at length to the entire class on a topic they may choose for themselves. Apparently, there is great variation in the narrative models, structures, and devices used across cultures. Children may experiment with many different types of narratives but it also happens that, because of their cultural background, they do not participate at all (Cazden, 2001). Cazden describes two striking examples of such non-participation. One was found by Philips (2001) in a study of the interaction patterns of native American children on a reservation in Oregon. Although encouraged by the teacher to talk in the classroom the children failed to do so. Apparently, they were unfamiliar with the social conditions for participation, as in the Indian community these conditions were lacking. Earlier, in 1982, Heath had found that black children did not participate in Appalachian schools, simply 'because they were not used to known-answer questions about the labels and attributes of objects and events' (Heath, 1982, in Cazden, 2001, p. 70). Also, from a pedagogic point of view, the educational value of sharing time depends very much on how it is handled by the teacher. Having pupils 'tell a story' in front of the classroom without proper feed forward and feedback is not liable to turn it into a strong learning experience.

The Initiation-Response-Feedback or I-R-F pattern has been investigated for decennia by dozens of researchers. Before it was called that way, Hoetker (1968) reported on classroom talk observations in nine junior high school English classes. Specific observations were made of the number of questions asked by the teacher in teacher-pupil talk. According to Hoetker the mean questioning rate per minute of substantive talk was one teacher question every 11,8 seconds. On first sight, by asking so many questions, teachers seem to involve their pupils in classroom exchange in quite a substantive way. But with the pupil response taking place within some fraction of that period (Hoetker, 1968) one cannot but question the quality of this involvement and of this kind of classroom talk.

Hoetker's findings were confirmed by Sinclair and Coulthard (1977) who analysed the question-response pattern of some 300 English secondary school teachers. They found that teachers mostly do not ask questions about things they do not know (and hope to learn from their pupils), they do so in the first place to find out if their pupils know the correct answer (Sinclair & Coulthard, 1977). In other words: the teacher knows the answer and is just checking whether his pupils do as well. Sinclair and Coulthard labeled this discourse pattern as I-R-F, or Initiation-Response-Feedback, e.g.:

Teacher:	What do you call animals like elephants, deer and bears?	**I**nitiation
Pupil:	Mammals?	**R**esponse
Teacher:	Right, mammals	**F**eedback

Figure 2. Example of the I-R-F model

In 1979 Mehan confirmed Sinclair and Coulthard's findings calling it I-R-E (Initiation-Response-Evaluation) and adding the thought that it is a form of communicative competence specific to schools, which has to be learnt but is not explicitly taught (Mehan, 1979).

Although the I-R-F pattern was identified and described nearly forty years ago, it is still omnipresent in classrooms today (Ellis, 2012; Hiebert & Stigler, 1999). The pattern has become almost universally accepted as the essential teaching exchange (Edwards & Westgate, 1994) or as an example of how power relations are supported by and played out in discursive practices in every classroom (Hicks, 1995-1996). Some researchers criticise what could be considered a mistaken belief that it encourages the student's participation (Lemke, 1990), others have found that dialogue structured in terms of evaluative I-R-F sequences might have an interactive authoritative character (Mortimer & Scott, 2003) that 'does not allow the synthesis/integration and exploration of ideas' (Scott, Mortimer & Aguiar, 2006, p. 605). Or, as Myhill and Dunkin (2005) point out, factual questions enable lessons to proceed more rapidly and help the teacher in maintaining order in the classroom. Because it does not allow exploring ideas or thoughts presented by students Van Lier (1996) criticises the traditional I-R-F pattern as not representing 'true joint construction of discourse' (Van Lier, 1996, p. 151). Rather, it is a means for the teacher to lead the lesson to a planned direction. Also, it enables him to hand out speaking turns so that he can control classroom interaction (Van Lier, 1996). The pattern leads pupils through a predetermined set of information and does little to encourage them to express their thinking (Cazden, 1988; Nystrand, 1997, in Mercer & Littleton, 2007). Moreover, while teachers are not always aware of their use of the I-R-F sequence, pupils are very much aware of it and they even act accordingly. Mehan (1979) found more student-initiated actions at the end of the school year than at the beginning. This seems to point out that pupils implicitly improve their 'I-R-F skills', but whether this adds to the quality of classroom talk, remains questionable. Galton (2002, in Coultas, 2012), for instance, found evidence that pupils develop strategies to avoid taking part in this pattern of talk, e.g. by giving the impression that they need more time to think. Lefstein (2008) even found that pupils deliberately refrained from answering a difficult open question, as they knew the teacher would eventually narrow the scope of it and ask an easier, closed question.

The I-R-F sequence seems to benefit teachers only

The rather negative qualitative assessment of the I-R-F pattern has also been confirmed by quantitative data. In 1970, based on findings in six different US school systems, Flanders and his research staff formulated the 'rule of two-thirds': they found that two thirds of the average lesson period in secondary education is devoted to listening and speaking ('someone is talking'), while two thirds of these activities are taken up by the teacher (Flanders, 1970). This means that on average, in a 50 minute lesson period, a teacher monopolises 20 minutes of all language activity. Completing the 'rule of two-thirds', Flanders (1970) found that some two-thirds of the teacher's language activity consists of expressing his own opinions or facts (lecturing), giving directions and criticising students, leaving little room for any form of dialogic teaching. In some classrooms the rule of two-thirds even appeared to be the 'rule of three-fourths or more' (Flanders, 1970, p. 179).

In a similar study in the US by Dillon (1988), pupils in elementary school were found to ask only one question per month over an average school year, whereas teachers asked 84 questions in one hour. According to Dillon, inhibiting factors were the teachers'need to cover the curriculum and possible discipline problems if the children were allowed more opportunities to talk. Based on review studies since 1912, Watson (1985) provides even more striking statistics, saying that teachers commonly ask 50.000 questions a year while pupils ask no more than ten questions per year each. Additionally, English, Hargreaves and Hislam (2002) found that during the 'literacy hour' in English primary schools the oral contributions of nine pupils out of ten consisted of three words or even less. Uninterrupted interactions of more than 25 seconds between the teacher and a pupil or group of pupils had declined after mandatory highly structured whole class activities were introduced (English et al., 2002; Topping & Trickey, 2014). Eventually, Mercer (1995) and Mercer and Littleton (2007) confirmed that opportunities for pupils to really discuss what they are learning in the classroom are rare. The ever-dominant I-R-F pattern, which seems inherent to the one dimensional view of language as knowledge transfer vehicle, hardly gives pupils the chance to use talk in order to learn.

Summarising, the I-R-F discourse pattern enables teachers to control the 'agenda', to maintain discipline in the classroom, to provide additional information and paraphrase a pupil's answer in a way that models for the other pupils how to phrase the statement in the academic jargon and make clear what they have to learn. The fact that teachers ask lots of questions in the classroom seems to suggest a very democratic and socratic behaviour, but as the I-R-F pattern remains a one-dimensional and restrictive way of communicating, it rather suggests the opposite. All these findings suggest that the use of language as an effective learning tool in the classrroom has long been – and still is – problematic.

The transmission model for learning

So why do teachers persistently use this I-R-F pattern? Is it indeed only a matter of controlling the classroom, checking understanding and modelling answers? Watson and Young (1986) believe it also has to do with teachers' (implicit) view on language in education, i.e. that of language as a transfer vehicle and not as a learning tool. As mentioned earlier, Barnes (1976) calls this view or practice the *transmission model*. Watson and Young are very sharp in their rejection of this model when they write: 'It seems that the majority of teachers lack faith in the capacity of their pupils to be active, constructive participants in their own learning. They feel that they must *tell* their pupils what to do, *interpret* new knowledge for them, *make explicit* any generalisation that can be drawn from the accounts of the experience being presented rather than structure the classroom so that pupils feel the need to develop their own accounts more fully' (Watson & Young, 1986, p. 129).

The next, logical question then is: where does this transmissional view come from? In 1977, Barnes and Todd found that teachers have a rather static view of knowledge. They reduce it to an objectively transferable object (Barnes and Todd, 1977; Barnes & Todd, 1995a). Relying on this view, many teachers fail to connect subject matter to their pupils' pre-existing knowledge. Consequently, pupils do not feel stimulated to engage in an active process of learning. On the contrary, they learn that learning is a passive, receptive process. Barnes and Todd found confirmation in the findings of Bliss, Askew and Macrae (1996) according to whom pupils are seldom encouraged to consider the lessons they attend as opportunities to explore their own ideas and opinions. However, this does not mean that teachers do not want to consider more effective ways to use language as a means to learn. Myhill and Dunkin (2005) acknowledged teachers' efforts towards a more reflective and collaborative teaching but found that many of their factual classroom questions were typical for the I-R-F pattern, and elicited thinking. About this inconsistency both researchers write: '[…] it is as though teachers want to open up pupils' thinking and reflection but cannot relinquish the control of discourse afforded by factual questions' (Myhill & Dunkin, 2005, p. 426). And though teachers may very well be willing to stimulate their pupils to learn through language, many do not know how to achieve this. Outside the UK and the US, some researchers have stressed the impact of culture on education, more specifically on the lack of dialogic tradition because of the emphasis on transmission as well (Nikolaidou, 2012; Rojas-Drummond et al., 2003).

Based on the large amount of research about the occurrence of the I-R-F pattern and the rather transmissional view on education which it reflects one would be inclined to think this is a universal phenomenon. It is not. A very important issue left untouched so far is the influence of educational tradition and culture on the place of classroom discourse in the curriculum and in teaching. This has been thoroughly investigated by Alexander (2001; see also 2004, 2008; Mercer & Hodgkinson, 2008). His evidence shows that the extent to and the way in which teachers implement talk in the classroom very much reflect the way they think people should relate to each other and that is not universal. Alexander investigated classroom talk and the broader context in which it fits in Russia, France, England, India and the United States of America. From his data he deduced three versions of human relations, or three relational values: individualism, community and collectivism. Individualism focuses on the self, personal rights and freedom of action. Community focuses on human interdependence, a principle which makes it imperative that people care, share and collaborate. Collectivism also focuses on human interdependence but its prime function is to serve the larger needs of society or the state. These three versions of human relations or values, Alexander found, ‘influence the dynamics and communicative relationships of classroom talk’ (Alexander, 2001, p. 96; 2004, 2008; see also Mercer & Hodgkinson, 2008).

Alexander’s data also laid bare six recurrent pedagogical values, one of which is *teaching as transmission*, which is very similar to Barnes’ transmission model, and another is *teaching as negotiation*, the practice of which comes close to Barnes’ interpretative model. Relational and pedagogical values seem connected to each other, as Alexander illustrates. In countries like France and Russia, for instance, knowledge transmission through mostly whole-class teaching is one of the more important educational goals, as community and collectivism are much more part of their culture (Alexander, 2008). Contrastively, in many American classrooms an antipathy towards transmissional teaching has been developed to such an extent that interaction has almost become a means to an end. ‘Pupils individually expressed their own mathematical meanings, say, but lacked a common language collectively to make sense of and evaluate them’ (Mercer & Hodgkinson, 2008, p. 99).

Summarising, the transmission model may be dominating the educational world, but it is not universal and, more importantly, it seems to serve functions which go back to cultural values we are not always aware of. Alexander’s findings may be less relevant for this dissertation, which concentrates on a specific type of talk for learning in one culture, but they have much practical value, as today’s classroom population is rapidly becoming more culturally diverse.

The interpretative model for learning

It has already been mentioned that a sound alternative for the transmission model is what Barnes calls the interpretative model (Barnes, 1976). Both have strong constructivist characteristics. In the interpretative model, the teacher takes into account the active nature of children’s learning and – as the cognitive psychologists already knew – stimulates them to connect new knowledge with what they already know. Preferred ways to do this, Barnes suggests, are exploratory discussions and writing in order to construct knowledge for themselves. Several researchers (Alexander, 2004; Mercer, 1995; Wegerif & Mercer, 1997b) would add to this that teachers and pupils jointly create knowledge and understanding. But, as Watson and Young (1986) explain, the move from transmissional teaching to interpretative teaching is not easy. Watson and Young advise a three step transition: a) encourage pupils to develop their answers ('make them talk into understanding', cf. Barnes, 1976; b) actively 'relax' the degree of conversational control and create a 'free speech' classroom, so that less pupils' responses have to be channeled through the teacher; and c) make a critical analysis of textbooks pupils have to use, as many of these rather seem to reinforce the negative aspects of the I-R-F pattern (Watson & Young, 1986).

According to Mercer (2010b), there are three good reasons why educational policy makers, teachers and educationalists should find speaking and listening in education important. The first is that for teachers, speaking and listening are about the most vital basic teaching tools. Therefore, they should use these resources as efficiently as possible. The second reason is that children and adolescents need to learn to communicate effectively. We should not assume that children learn to speak effectively and strategically without any guidance. The third reason is that the degree to which children are involved in learning conversations (classroom learning conversations as well as conversations during duo and group work) has a positive effect on their thinking and learning. The quantity and quality of speech that children experience in their first years at school, predict very well as how well they will perform in secondary education (Hart & Risley, 1995). According to Brown (2016) this statement seems to have lost none of its topicality value, because 'despite the benefits students gain through argumentation, research has indicated that argumentation skills and practices are lacking among adolescents' (Brown, 2016, referring to Zhang & Lu, 2014, and to Jonassen & Kim, 2010).

By using language and experiencing how others use it, pupils learn to describe the world around them, to handle their own experiences and to get things done. They use language to think, either alone or together. This leads us to the focus of this dissertation: pupil-pupil talk. Compared to 20 years ago we now know a lot more about the importance of spoken dialogue to learn and develop speaking and reasoning skills (Mercer, 2010b). Unfortunately, there seems a need in education to assume that, once pupils move on to secondary school, they know everything a person has to know about speaking and listening there is to learn. Teachers ask them to work together in groups and assume that they can do so without problems. This is a misconception, as Setati et al. (2002) put it: '[…] learning from talk is significantly limited if it is not supported or complemented by strategies for learning to talk, i.e. learning subject-specific formal or educated discourses' (Setati et al., 2002, p. 147).

Indeed, many pupils have little or no experience in conducting rational discussions. It is a skill they are not born with and only few are fully raised with it. When teachers realise how (classroom) learning conversations (can) go, they may learn the impact of talk on the degree to which their pupils actually participate in the learning process. It is clear that the I-R-F pattern does not make pupils learn through talk. The interaction needed between teacher and pupils is not an I-R-F pattern but an IDRF-pattern (Wegerif in Mercer & Littleton, 2007), where D stands for 'Discussion'. In the same respect, there must be room for discussion among pupils during class time or, even better, during collaborative work. By consequence, as Bullen et al. (2002) elaborate, ‘the classroom provides a framework for both the acquisition of knowledge for the individual and for the development of social interaction with others’ (Bullen et al., 2002, p. 206) and ‘collaborative skills are embedded within a socio-cultural context and thus are fundamental for both learning in school and in adult life’ (Bullen et al., 2002, p. 203).

Together with IDRF exchanges between teacher and pupils, active learning through partner or group work is imperative for learning through talk. When pupils are allowed to work in small groups teacher language activity decreases significantly. The question that remains is: how effective can talk be in a collaborative work form like group work? This is the subject of the next section.

3. Language and learning during group work

As collaborative learning is one of the principles of social constructivism, it is obvious various work forms have been developed that stimulate pupils to work together. Group work is one of the work forms that encourage the exploration of ideas (Barnes & Todd, 1995b). In group talk pupils can risk hesitation, confusion and rejection of their ideas by their peers. Moreover, when pupils feel secure, they can think aloud and reshape and interpret ideas (Enghag, Gustafsson & Jonsson, 2007). Pupils have to develop listening and speaking skills and receive significant opportunities to practice. Children develop their language skills by interacting with one another and as they participate in conversation more actively, their language development improves.

Group work versus working in groups

There is a huge difference between group work and working in groups. When pupils are assigned to do the same work in dyads or small groups, for which they do not really need one another, Hoogeveen and Winkels (2008) call it ‘working in groups’. Their findings resemble those of Dawes et al. (1992) who suggest that without sufficient preparation, true collaboration easily becomes parallel working, making group work ‘a wasted educational opportunity’ (Mercer, 2010b). This is often the case when the assignment is too easy or when collaborative skills are not stressed upon. Real group work is more than that. It implies that pupils learn to work together in an efficient way in order to achieve certain (learning) goals. Slavin calls this *cooperative learning*, which he defines as ‘instructional programmes in which students work in small groups to help one another master academic content’ (Slavin, 1996, p. 200). Coffey elaborates on this definition, rephrasing it as: ‘an instructional method in which students work together in small, heterogeneous groups to complete a problem, project, or other instructional goal, while teachers act as guides or facilitators’, to which she adds: ‘This method works to reinforce a student’s own learning as well as the learning of his or her fellow group members’ (Coffey, 2016, p. 1).

Requirements of efficient group work

What requirements must efficient group work meet? Many publications focus on organisation and structure: clear goals and instructions, the right size and composition of groups, shared and individual responsibilities, specific kinds of group work, collaborative strategies and systematic assessment (Kagan, 2014; T'Sas, 2013; Hoogeveen & Winkels, 2008; Paelman, 2004; Slavin, 1996). School practice as well as research have demonstrated that teachers do not find it easy to meet all these requirements (Panitz, 1997). If they fail to do so, group work risks becoming chaotic, the teacher may lose control and the added value for learning diminishes. For many teachers this is the reason why they refrain from organising group work whatsoever (Miller, Trimbur & Wilkes, 1994; Panitz, 1997). Especially inexperienced teachers and student teachers are vulnerable. Not only do they have a lot to learn about organising group work themselves, often they have to battle against the expectations of their pupils, which have been formed during tuition by other, more tradition-minded teachers (Coultas, 2012; Fisher, 2011).

Group work and types of discourse

One essential aspect of group work that has not been mentioned so far, is language, i.e. the quality of group talk itself. During various experiments in British primary and secondary schools Neil Mercer and his research team made recordings of partner work and group work of pupils and analysed the conversations that took place there. Most of those group conversations seem rather unproductive. Based on their observations Mercer (1995) distinguishes two kinds of educationally unproductive talk: *disputational talk* and *cumulative talk.*

Disputational talk is marked by individualism, competition, disagreement and the lack of proper argumentation. Mercer (1996) gives the following example:

Lester:	1, 2, 3, 4, 5 (***counting grid squares on the screen with his finger, before he takes his turn***)
Sean:	1. It's there.
Lester:	So it has got to be . . .
Sean:	5, 4 (***suggesting a set of co-ordinates***)
Lester:	(***ignoring Sean***) . . . 4, 3. No, we have had 4, 3.
Sean:	4, 5. No, 4, 4.
Lester:	4, 3 (***presses keys for his turn***). What! (***he fails to find the elephant***). That's easy, I know where it is, opposite.
Sean:	(***sits silently for a while, looking at the screen.***)
Lester:	I can do it.
Sean:	(***Still staring at the screen***) No, not up, down.
Lester:	It can't be.
Sean:	It can.
Lester:	I know where it is.
Sean:	(***eventually takes his turn, but fails to find the elephant.***)
Lester:	I told you it weren't over there. (***He then takes his turn, without success***)
Sean:	Eh heh heh heh. (***laughing gleefully***)

Figure 3. Illustrative fragment of disputational talk (Mercer, 1996, p. 366)

In this sequence, as Mercer explains, two 10-year-old boys, Sean and Lester, are at the computer, doing a puzzle. They have to find an elephant by keying in co-ordinates. Though the boys appear to be working enthusiastically their conversation is full of short assertions, rebuttals or comments which are not constructive. They take alternate turns, as may be expected in a collaborative conversation 'but then so do opponents in tennis' (Mercer, 1996, p. 366). The atmosphere is very competitive and the boys are far from being really collaborative.

Cumulative talk is marked by the quick and uncritical piling up of ideas, no real argumentation, superficial consensus and often the urge to 'get the job done' as fast as possible. The following transcript illustrates this:

Katie:	Okay, so right then. What shall we write?
Anne:	We can have something like those autograph columns and things like that and items, messages
Katie:	Inside these covers (*pause 3+ secs*). Our fun filled
Anne:	That's it!
Katie:	Something
Anne:	Something like that!
Katie:	Yeah
Anne:	Inside this fabulous fun filled covers are – how can we have a fun filled cover? Let me try
Katie:	Inside these (*pause 3+ secs*)
Anne:	Hah huh (*laughs*)
Anne:	You sound happy on this. Fantabuloso (*laughs*)
Katie:	Inside these inside these fant, inside these fun-filled, no inside these covers these fantastic these brilliant
Anne:	Brilliant
Katie:	Is it brilliant?
Anne:	No
Katie:	No Fantast fantabuloso shall we put that?
Anne:	Yeah (*inaudible*) fantabluloso
Katie:	Fan-tab-u-lo-so
Anne:	Loso. Fantabuloso.
Katie:	Fantabuloso oso
Anne:	Fantabuloso ho!

Figure 4. Illustrative fragment of cumulative talk (Mercer, 1996, p. 367)

In this sequence, Katie and Anne ask each other questions about a text, make suggestions and offer reasons for the decisions they take. Collaboration seems fine, but not really productive: 'They confirm and validate each other's statements, explicitly ('That's it') or implicitly by repeating them ('Inside these ...'). They are not only constructing their text together, they are constructing a joint understanding of what the text should be like. […] There is no real disagreement: they do not challenge each other's suggestions, and do not seem to feel the need to justify opinions or explain their reasons' (Mercer, 1996, p. 367).

The effects of both types of talk are mostly negative: a lot of noise, loss of time, low quality of group work and unfair distribution of responsibilities (Mercer, 2010b). These may be extra reasons why teachers are hesitant to do or reject group work, the unfortunate side effect being that pupils miss powerful learning opportunities.

To keep these learning opportunities intact, pupils have to learn how to talk and work together. Instead of competitive or cumulative talk, Mercer (1995), following Barnes (1976), advocates to organise group work in such a way that pupils conduct *exploratory talk*. This type of talk has the following characteristics: pupils listen attentively to each other; they ask each other questions; they share relevant information with each other; any idea or opinion can be questioned; pupils explain why they are criticising or questioning someone's ideas and opinions; pupils build on what has already been said; everyone is encouraged to participate; pupils respect each other's opinion and idea; there is an atmosphere of mutual confidence in the group; the feeling is that the pupils jointly work towards a goal; and the group strives for consensus and looks for arguments (Mercer & Littleton, 2007).

Mercer illustrates exploratory talk with the following transcript:

Diana:	Let's discuss it. Which one shall we go for?
All:	(*inaudible—reading from instructions*)
Peter:	1 2 3 or 4 (*reading out the number of options available*). Well we've got no other chance of getting more money because
Adrian:	And there's a monastery
Diana:	And if we take number 2 there's that (*inaudible*)
Peter:	Yeh but because the huts will be guarded
All:	Yeh
Adrian:	And that will probably be guarded
Diana:	It's surrounded by trees
Peter:	Yeh
Adrian:	And there's a rock guarding us there
Peter:	Yes there's some rocks there. So I think, I think it should be 1
Adrian:	Because the monastery might be unguarded
Diana:	Yes 1
Adrian:	1 yeh
Peter:	Yeh but what about 2? That, it might be not guarded. Just because there's huts there, it doesn't mean it's not guarded, does it? What do you think?
Diana:	Yes it doesn't mean it's not. It doesn't mean to say it's *not* guarded, does it? It may well be guarded. I think we should go for number 1 because I'm pretty sure it's not guarded
Adrian:	Yeh
Peter:	OK, yes number 1 (*he keys in 1 on keyboard*). No (*computer responds inappropriately*)
Adrian:	You have to use them numbers (*he points to the number keys on right of board, and Peter uses them to obtain the right result. Adrian begins to read from screen display*) 'You have chosen to raid area 1'.

Figure 5. Illustrative fragment of exploratory talk (Mercer, 1996, p. 368)

These children (aged 9 and 10) are using a computer program called Viking England. They take on the active roles of Viking raiders planning an invasion of the English coast. They are discussing which target they would raid: a monastery, a village of huts, a castle or a harbor. In this sequence the children are very much on task, 'asking each other questions, commenting and making suggestions. They discuss the various options, and also remind each other of relevant information. They are using talk to share information and plan together. They discuss and evaluate possible courses of action and make joint decisions. […] Reasoning is essentially interactive, not really reducible to the form and content of individual statements, but more to do with how the discourse as a whole represents a social, shared thought process' (Mercer, 1996, p. 368).

We will discuss exploratory talk extensively in Chapter 3, but before that, we wish to mention the meta-analysis about the effects of a number of variables on collaborative learning, made by Janssen, Kirschner, Erkens, Kirschner and Paas (2010). According to Janssen et al., since the beginning of the 20^{th} century more than 1200 studies have been carried out to examine the effects of collaborative learning on cognitive, meta-cognitive, affective-emotional and social output. However, most of these studies do not explain why some groups of pupils work well together and others do not. To Janssen et al. (2010), this is because one important variable tends to be disgarded, i.e. the interaction between pupils while collaborating in a group. After thorough selection based on their research questions these authors analysed 124 studies which investigated either the effect of didactics on interactive processes, or the effects of interactive processes on learning outcomes or both.

With 'didactics' Janssen et al. (2010) mean everything the teacher can 'manipulate' when organising collaborative learning: group formation (homogeneous or heterogeneous for gender, ethnicity, ability); task (divergent or convergent, face-to-face or computer-mediated, the extent to which it creates interdependency); tools, support and scaffolds; and preparation (collaborative strategies, language strategies). Interactive processes are about a) the function of the interaction, i.e. its contribution to knowledge and insight construction and problem solving, and its contribution to the regulation and maintenance of the group process, and b) the way the interaction runs (high or low elaboration, talk about task/content). We have summarised the findings of Janssen et al. in table 1 and 2.

Table 1

Effect sizes of didactic factors on interactive processes, participation, elaboration, task focus and group process regulation, based on Janssen et al. (2010)

Independent variables			**Effect on**			
			Dependent variables		**Effect size**	**Measured effects**
Didactics	heterogenuous group formation	ability	interactive processes	in general	24	-0.19
		gender		in general	6	+0.13
		other		in general	13	+0.25
	interdependency	high	interactive processes	off-task participation	31	-0.32
		high		high elaboration	31	+0.44
		high		low elaboration	31	+0.38
	communicative medium		interactive processes	In general	6	+0.07
	tools, support and scaffolds		interactive processes	on-task participation	44	+0.56
				high elaboration	44	+0.38
				low elaboraton	44	+0.23
				group process regulation	44	+0.25
	preparing pupils for interactive processes		interactive processes	on-task participation	197	+0.39
				off-task participation	197	-0.25
				high elaboration	197	+0.35
				talk about task/content	197	+0.15
				group process regulation	197	+0.20

Table 2

Effect sizes of interactive processes on learning outcomes, based on Jansen et al. (2010)

Independent variables		**Effect on**			
		Dependent variables		**Effect size**	**Measured effects**
Interactive processes	high elaboration	learning outcomes	in general	174	+0.17
			high elaborative utterances	174	+0.34
			high elaborative support	174	+0.21
			high elaborative response	174	+0.14
	low elaboration	learning outcomes	in general	174	-0.08
	discussing task/content		in general	174	+0.39
	group process regulation		in general	174	---
	on-task participation		in general	174	+0.22
	off-task participation		in general	174	-0.17

Janssen et al. (2010) show that many factors positively influence interactive processes on the one hand and a number of aspects of interactive processes positively influence learning outcomes on the other. In order to determine how decisive or how important these factors are, Janssen et al. suggest using Hattie's (2009) cut-off point of +.40 effect size. If we use this standard as well, then tools, support and scaffolds are the most influential variable in didactics (+.56). Also influential is interdepency (+.44). Variables which come very close to the cut-off point are: a) for didactics: preparation for interactive processes on ‘on task’ participation (+ .39) and on high elaboration (+.35), and interdepency on low elaboration (+.38); and b) for interactive processes: talk about task/content (+.39).

Based on these findings it may be concluded that preparing children to work together, in one way or another, appears to be very important to make collaborative work successful in the first place. Second, the use of the tools, scaffolds and support is almost equally important. A third factor that makes a significant difference is interdependency. The next chapter will reveal that interventions to promote exploratory talk correspond with these three facors.

4. Summary

In the first section of this theoretical framework we examined the growing role of language in modern learning theories. Behaviourist thinking considered learning as a change of behaviour, induced by a stimulus-response pattern in which language is a mere transfer vehicle. Cognitivists and - a fortiori - cognitive constructivists saw language as a framework for constructing and integrating knowledge through active learning. Social constructivist Lev Vygotsky put language in a central mediating position between social and individual learning: to him cognitive development is a process that originates in and must be explained by products of social interaction, with the teacher or peer in the role of scaffolder and the learner employing metacognitive strategies to enhance his learning. Recent findings in neuroscience confirm the importance of active learning. Neuropedagogists stress the fact that using language involves higher-order thinking which in turn exhibits much learning potential. Higher-order thinking, then, is reflected in higher-order talk of the kind that becomes visible in exploratory talk.

In the second section we made up a status quaestionis about the factual and preferred use of language for learning in the classroom. Decades of research show that language has been and is being used to transfer rather than to jointly construct knowledge at school. Almost worldwide, teachers, whether consciously or not, use the (ever) dominant I-R-F pattern (Initiation-Response-Feedback) in their classroom exchanges, leaving little room for any form of dialogic teaching. In this transmission model, pupils receive little stimulus to express their thinking and share it with the teacher, let alone with their peers. The (social constructivist) alternative for this transmission model of education is the interpretative model. Here the teacher acknowledges the active nature of children's learning and stimulates it by implementing e.g. collaborative learning activities. As speaking and listening skills predict how young school children will perform in secondary school, it is necessary to start their dialogic learning process early enough, preferably in the first years of primary school.

The third section zoomed in on the effective use of language in group work, which is considered a strong pedagogic tool to valorise learning potential in the classroom. However, it does not suffice to split pupils up in groups and expect a smooth learning process. Group work has to be well-organised, well-structured, pupils need clear goal instructions and tasks must be designed in such a way as to create interdependency. But apart from these conditions, conversation itself requires attention. Without proper guidance and preparation, pupils may lose opportunities for learning by degrading their conversations to disputational or cumulative talk. Exploratory talk, then, has the most learning potential but in order to master it, pupils must learn how to use it.

This leads us to the next chapter, where we will explore and examine the definition of exploratory talk, how it is measured and which effects it generates.

Chapter 3: Study 1: Literature study

In this dissertation we want to get a firm grip on the concept of exploratory talk, implement it in a classroom experiment and measure its effects on pupils' use of exploratory talk and problem solving capabilities. This chapter focuses entirely on the concept of exploratory talk and answers RQ 1: What is exploratory talk, how is it measured and which effects does it have?

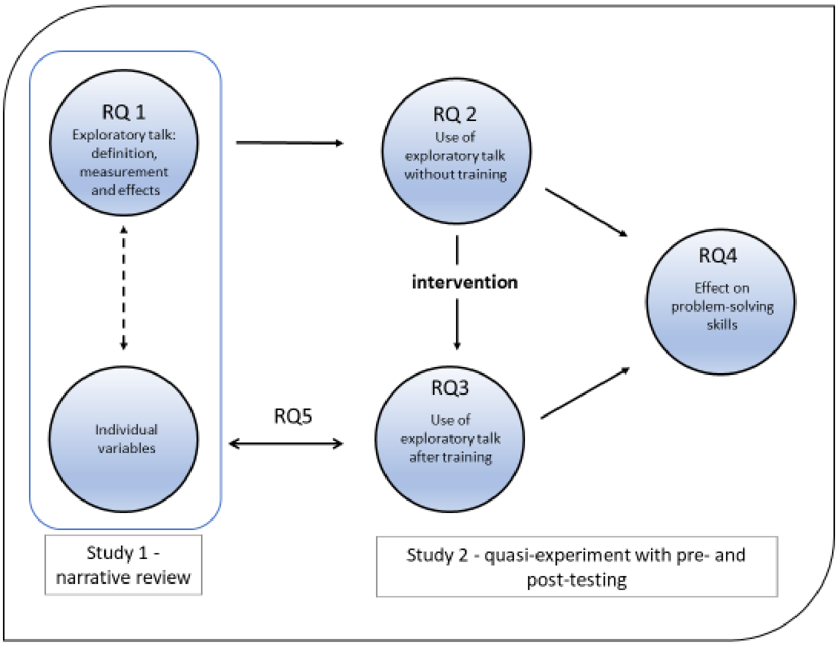


Figure 6. Global Research Design

As explained in Chapter 1, 'Problem, research questions and research design, the first part of this dissertation is the narrative review, which answers RQ 1 by determining the definition, measurement and effects of exploratory talk. We need this information to proceed to the next step, i.e. to construct and carry out the quasi-experiment (Study 2) which enables us to answer the other research questions. The extensive literature search will also shed more light on the importance of certain individual variables in order to eventually answer RQ 5.

In order to define the concept of exploratory talk, to describe how it is measured and to give an overview of possible effects of this type of talk, a database search was conducted. The result of this search and analysis is a narrative review of 115 articles (international, peer reviewed). We chose the format of a narrative review because it enables us to establish a holistic view on the topic, leaving room for input from existing theories and reflections based on personal experience and shared educational knowledge.

As the concept of exploratory talk goes back a long way (Barnes, 1976) and has been studied extensively, we expect to find a clear-cut definition and description of its characteristics, as well as a methodological consensus about the measurement of this type of talk. We also expect to find an overview of proven learning effects.

1. Exploratory talk: 40 years of research

As mentioned in the previous chapter, the quantity of talk in the classroom can be augmented through methods for active learning, but what about the quality? The first to mention the educational value of this specific way of talking, i.e. *exploratory talk*, was Douglas Barnes (1976). In the mid-nineties, following the growing implementation of the constructivist view on learning in school curricula, and considering the fact that interaction is an important aspect of the constructivist view on learning, the concept of exploratory talk was further explored and defined by a line of British researchers who largely focused on pupil-pupil talk (e.g. Fisher, 1993; Mercer, 1996; Wegerif & Mercer, 1997a). During classroom experiments pupils of different ages and backgrounds - usually in primary education - were taught to use exploratory talk during collaborative work. Teachers were taught similar skills in order to raise the learning quality of classroom talk or they were directly involved by teaching their pupils how to use exploratory talk.

As pre- and post-tests repeatedly demonstrated its educational value (for an overview, see Mercer and Hodgkinson, 2008), interest in exploratory talk for learning has kept on growing: classroom experiments have been and are being replicated, with researchers using familiar but nevertheless varying descriptions of exploratory talk and/or refining and elaborating on the concept. Researchers have also brought in or focused upon specific variables, using an increasing number of different analytical models, and leaving critical issues and additional questions to be tackled. This widening conceptual and methodological approach towards exploratory talk may be reflecting multiple interpretations or even doubts about the definition, the characteristics and the optimal analytical tools and strategies concerning its measurement. Therefore, in this review we want to make a round-up of the research process as it has developed since 1976 and see where it has lead us, both conceptually and methodologically. Specifically, we want to determine whether the early descriptions of exploratory talk as a type of pupil-pupil talk still stand. We want to determine its characteristics and the analytical methods used to label classroom conversation as exploratory (or not). In short, with this review we want to answer the first global research question of this dissertation:

RQ 1 How is exploratory talk defined, how is it measured and which effects does it have?

Subdivisions of this research questions are:

RQ 1.1 How is exploratory talk defined?

a. In what way has the concept of exploratory talk evolved in terms of its definition and characteristics (concept growth)?
b. Have any (near) synonyms and antonyms been introduced?
c. Have any elaborating and deepening/refining terms been proposed?
d. Has exploratory talk been embedded in any overarching terms?
e. Which definition of exploratory talk is currently being used?
f. What is the contextual use/relevance of exploratory talk?

The answer to RQ 1.1 will be the onset to answer RQ 1.2 and RQ 1.3:

RQ 1.2 How is exploratory talk measured?

a. What kind of studies have been undertaken (theoretical, experimental …)?
b. Which research methods have been used (quantitative, qualitative, mixed methods)?

RQ 1.3 Which effects does the use of exploratory talk generate?

a. Which characteristics of exploratory talk and variables influencing it have been focused upon?

2. Methodology

In order to answer the research questions, a literature search was conducted. Six electronic databases were searched: Communication & Mass Media, Education-line, ERIC, JSTOR, Linguistics and Language Behaviour Abstracts (LLBA) and Web of Science. Google Scholar was also searched. The primary literature search was limited to the years 1976-2016, as in 1976 the first article was published in which the term 'exploratory talk' was used in the context of educational research (Barnes, 1976). In order to be included in this review, all manuscripts had to be peer reviewed. As such, the main search term 'exploratory talk' yielded 49 articles in ERIC, 58 in Web of Science, 13 in JSTOR, 18 in LLBA, 12 in Communication & Mass Media, 127 in Education-line and 84 search entries in Google Scholar which mentioned 'exploratory talk' in the title of the articles ('exploratory talk' searched as bare text in Google Scholar rendered 4500 entries). Although the search results showed some overlap, browsing the abstracts also revealed that a number of articles were not relevant for this review, either because the term ‘exploratory talk’ was not used in an educational context (e.g. politics, business) or because it was used as a descriptive term rather than a type of talk. We therefore decided to refine our search.

In a second step, we combined the main search term ‘exploratory talk’ with additional terms which were collected and categorised via peer debriefing (Figg, Wenrick, Youker, Heilman & Schneider, 2009). Four concepts were defined: education (concept 1), language (concept 2), context (concept 3) and theory (concept 4). We used education and its subterms as the first concept, as these were most likely to rule out any non-educational context in which the main term would be used. Language, context and theory were added as second, third and fourth concept to determine the main focus of each article. In order to be reviewed, each article had to contain the main term, one or more (sub)terms of concept 1 and (sub)term(s) of at least one of the other three concepts. An overview of these search terms is presented in Table 3. The combination of search terms resulted in a selection of 121 articles.

Table 3

Search terms hierarchy applied during literature search

Main concept	Combined with	And
exploratory + talk	Concept 1 : education	
	education	(none)
	classroom	(none)
	groups	(none)
		pupils
		students
	learning	teacher
		(none)
		collaborative
	teaching	interaction
		(none)
	Concept 2: language	
	language	(none)
		(linguistic) markers
		discourse/discursive
		reasoning
	Concept 3: context	
	problem solving	(none)
	attitude	(none)
	Concept 4: theory	
	Piaget	(none)
	Vygotsky	(none)
	(social) constructivism	(none)

Third, after ruling out doubles, 105 manuscripts were identified for further reading. In addition, the reference lists of these manuscripts were explored in order to look for other relevant manuscripts, i.e. articles being referred to by at least three different researchers or research teams and which did not appear in the databases we searched. These were mostly older manuscripts written by early researchers ('pioneers') such as Sinclair and Coulthard (1977), Barnes (1976), Barnes and Todd (1977). This way, 10 articles were added for analysis, totalling 115 articles for reading. From the 115 articles included in this review study, 63 enabled us to answer RQ 1.1 and 88 to answer RQ 1.2 and RQ 1.3.

Fourth, all selected manuscripts were read thoroughly in order to find patterns in the results. The nature of these search patterns was dictated by pre-defined themes based on the research questions. Article excerpts thought useful to answer the research questions were coded into these themes, using concordancer software, i.e. NVivo 10 (Bazeley, 2013). This software makes it possible to do an automatic search for words or word phrases in multiple documents. The following pre-defined coding scheme was applied:

Table 4

Coding scheme for the narrative review

Main item	Subitems	Including
Exploratory talk as a concept (RQ 1.1)	a. Concept growth	
	b. Defining exploratory talk	i. Definitions ii. Synonyms and antonyms iii. Elaborating/refining terms iv. Overarching terms
	c. Contextual use/relevance	
Measurement of exploratory talk (RQ 1.2)	a. Type of studies	
	b. Methods	i. Qualitative ii. Quantitative iii. Mixed methods
Effects of exploratory talk (RQ 1.3)	a. Variables and characteristics	

Finally, the themes were further explored in the manuscripts and incorporated into the narrative review. A narrative review summarises different primary studies which enables the researcher to draw conclusions into a holistic interpretation. This interpretation is merged with the reviewers' own experience, existing theories and models (Kugley, Wade, Thomas, Mahood, Jørgensen, Hammerstrøm & Sathe, 2016; Kirkevold, 1997). One of the strengths of a narrative review is its proposal to comprehend the diversities and pluralities of understanding around scholarly research topics and the opportunity to speak with self-knowledge, reflective practice and acknowledgment for shared educational phenomena (Jones, 2004). In short, we believe a narrative review brings the development of ideas to a deeper level and allows us to provide qualitative descriptions of the findings from literature (Dochy, Segers & Buehl, 1999).

3. Results (RQ 1)

Our search results suggest that research on exploratory talk was rather scarce in the years immediately after the introduction of the concept in 1976, but since 1986 it has increased over the years. Figure 7 shows the chronological distribution of the 115 articles discussed in this review. Growth manifestingly accelerated at the turn of the century, providing a maximum of 16 articles in 2010. Also, the topic was increasingly investigated outside the UK in the same period.

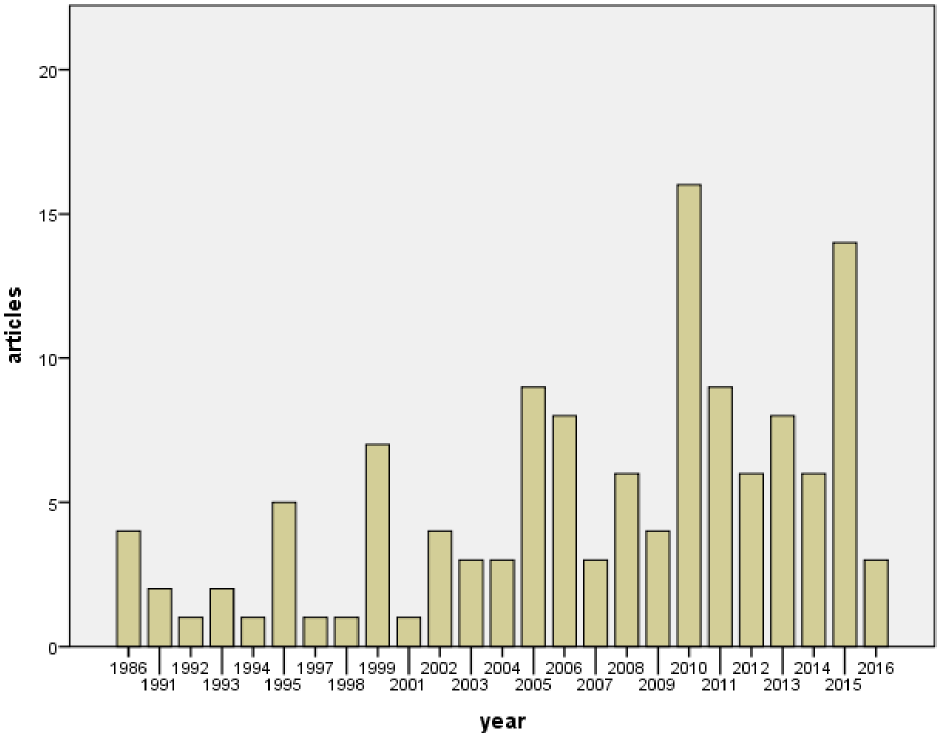


Figure 7. Distribution of the number of articles 1986-2016

In the following sections we will discuss what these studies told us about the definitions of exploratory talk (3.1), the ways it is measured (3.2) and its effects (3.3).

3.1 Exploratory talk as a concept (RQ 1.1)

As mentioned before, 63 articles enabled us to answer the question how exploratory talk is defined. We will now describe our findings based on the selected themes (nodes) (see Table 4). After briefly describing the concept growth in the first section, we will discuss the various definitions, (near) synonyms and antonyms, refinements and elaborations, as well as overarching terms of 'exploratory talk' in section two. In the third section, the contextual use and relevance of exploratory talk are discussed, meaning within which contexts exploratory talk is studied and why it is important to do so. We will end this paragraph by drawing conclusions and raising some issues for discussion.

3.1.1 Concept growth: description becomes definition

In this section we will discuss how the concept of exploratory talk came into existence. We will return briefly to Vygotsky and then make a crossover to a line of early Anglo-Saxon researchers, like Bruner, Britton and Barnes, who studied classroom talk in the 1970s.

As mentioned before, the databases show little research on exploratory talk between 1976 and 1986, and until 1995 the concept 'exploratory talk' remained descriptive, suggesting that it may not have been considered a concept or part of taxonomy of talk. We argue that Barnes' (1976) and Mercer's (1995) description and definition of exploratory talk are not only similar, until now they have not been refuted and remain the starting point for much research on this topic.

In *Thought and Language,* which was originally published in 1934 in Russian as *Thinking and Speech*, Vygotsky (1978) highlights the importance of language for learning suggesting it helps us to develop new ways of thinking. The greatest stimulus for this development comes from the interaction between a learner and his 'teacher' or anybody else who knows more about a subject and has the ability and willingness to support the learner in his learning. Vygotsky considers knowledge construction primarily as a social process: in our learning, we are influenced and stimulated by others. Language, then, has three major functions which work together: as a cognitive tool we use it to process knowledge; as a social or cultural tool we use it to share knowledge and as a pedagogic tool we use it to provide intellectual guidance to another (Mercer et al., 1999). Simply put, we use talk to develop our thinking and make sense to our experiences. It is only after this social process that we internalise and individualise what we have learnt collectively.

Bakhtin, Vygotsky's contemporary theoretician, developed similar thoughts about the role of language in learning, using the word 'dialogue'. To Bakhtin (2010), dialogue - in its dialectic sense - is essential when it comes to handling knowledge in an educational discourse and in learning. Vygotsky's and Bakhtin's theories were gratefully adopted by constructivist theoreticians, as, according to (social) constructivism, learning is only effective when students can play an active part in their learnings (Dewey, 1933), especially through exploratory forms of talk (Hardman in Mercer & Hodgkinson, 2008).

Renewed interest in Vygotsky's theory and its consequences for the link between language and learning has stimulated research thoroughly over the last four decades (for an overview, see McConaghy, 2014). They were stimulated by pioneers like Bruner, Britton and Barnes. For Bruner (1960), the meaning of ourselves and of our worlds is created through our own mental activity and mediated through cultural and social interaction. And so, Britton (1970) argues, we use talk to organise and reorganise or modify our own interpretations, while we 'greatly affect each other's representation' at the same time (Britton, 1970, p.6).

Building on a constructivist view of the nature or learning, Barnes (1976) stated that learning can never be a passive process. Whatever method or strategy a teacher uses, the pupils have to do the learning. As the essence of learning is connecting new knowledge to existing knowledge in an attempt to make sense of the world, pupils continually have to actively construct their new way of understanding. One of the readiest and most flexible tools to do so is language, i.c. talk, but not all kinds of talking have the potential to improve our understanding (Barnes, 1976). Barnes distinguishes between two kinds of talk: *presentational talk* and *exploratory talk*. Only in exploratory talk does the speaker concentrate on sorting out his thoughts and try to actively construct knowledge (Barnes, 1976). Barnes and Todd's (1977) research adds that pupils need ownership of knowledge if they are to discuss it openly and learn from one another. In other words, when working in small groups, with the teacher being largely absent from discussion, they will be more willing to risk rejection of their ideas by the other group members.

From 1990, focusing on talk while pupils are using computers in the classroom, researchers picked up Barnes' trail and started analysing pupil-pupil talk during problem solving tasks in pairs (Fisher, 1993; Galton & Williamson, 1992; Kruger, 1993; Phillips, 1992). Among many, Galton and Williamson (1992) argue that pupils must be taught how to collaborate if they are to do so successfully. Also, special attention must go to the language they use. Or, as Mercer (2010b) states: children are not born with the skills to talk effectively together or to develop specific dialogic strategies for thinking collectively. As Phillips (1992) acknowledges, the talking itself must have certain qualities in order to facilitate learning. Bluntly talking towards a consensus is bound to be less valuable for learning than exchanging arguments and counter-arguments in order to explore hypotheses. Exploratory arguments as a process of finding an answer or conclusion were found to be effective in fostering students' critical thinking and cognitive development (Phillips, 1992). Also, according to Kruger (1993), learning is linked to the quality of dialogue, particularly the amount of *transactive reasoning.* Light (1991) found that using language to make plans explicit, to make decisions and to interpret feedback seems to facilitate problem solving and promote understanding (Light, 1991).

Dawes, Fisher and Mercer (1992; see also Fisher, 1993), then, were the first to put children's talk during collaborative work into different categories, suggesting a taxonomy of *cumulative*, *disputational* and *exploratory talk*. Their claim is that only exploratory talk supports learning – or joint knowledge construction, a claim which is supported by a number of studies, see e.g.

Wegerif and Mercer (1997a). If talking is to lead to knowledge construction, partners must present ideas and give arguments to sustain them, when solving problems, issues must be jointly analysed and the discussion must lead to joint reasoning and decision making. In other words, the discussion must be exploratory (Mercer, 1996). In 1996, Mercer defined exploratory talk as a critical but constructive process of collective reasoning marked by specific attitudinal and linguistic characteristics (see the next paragraph for the full definition). Though there have been disagreements about how to measure exploratory talk and which characteristics to focus on (which is the subject of RQ 2), the definition itself has since then not been challenged substantially.

3.1.2 Defining exploratory talk

In this section we will discuss the definitions of exploratory talk and the extent to which the concept has been paraphrased, refined or deepened and embedded in larger concepts.

3.1.2.1 Definitions

Presentational and exploratory talk (Barnes)

The concept of 'exploratory talk' in an educational context was introduced by Douglas Barnes (1976). Based on a series of studies carried out in British primary schools Barnes distinguished, as mentioned earlier, between two types of classroom talk according to their relationship with learning: *presentational talk* and *exploratory talk* (between pupils only). Both contribute to learning, Barnes assures, but in the classroom each has a different place in the sequence of lessons. Presentational talk (also called 'final draft' talk) means that learners formulate a restrictive answer, mostly of the kind expected by the teacher. It is also referred to as 'Initiation, Response, Feedback' (Sinclair & Coulthard, 1977), a type of whole class dialogue in which the teacher focuses on understanding and correcting answers (cf. supra). In exploratory talk, however, pupils concentrate on sorting out their thoughts and try to actively construct knowledge (Barnes, 1976). Therefore, Barnes suggests that exploratory talk should come before presentational talk, as children need time to organise their thoughts.

Exploratory talk, Barnes explains, 'is usually marked by frequent hesitations, rephrasings, false starts and changes of direction. […] it is one means by which the assimilation and accommodation of new knowledge to the old is carried out […] the children not only formulate hypotheses, but are compelled to evaluate them for themselves' (Barnes, 1976, p. 28-29) and 'In exploratory talk […] the learner himself takes responsibility for the adequacy of his thinking' (Barnes, 1976, p. 113). To this he added, in 2008, a description which is since then being quoted more often than the one in Barnes (1976): 'Exploratory talk often occurs when peers collaborate in a task, when they wish to talk it over in a tentative manner, considering and rearranging their ideas. The talk is often, but not always hesitant, containing uncompleted or inexplicit utterances as the students try to formulate new understandings; exploratory talk enables students to represent to themselves what they currently understand and then, if necessary to criticise and change it. ' (Barnes in Mercer & Hodgkinson, 2008, p. 50). In this respect exploratory talk does not provide new information, '[…] rather students are able to make sense of something by sharing knowledge, explaining options, and examining ideas critically as they are being held publicly accountable' (Barnes in Mercer & Hodgkinson, 2008, p. 50). In other words, exploratory talk encourages thinking *par excellence* and is therefore the preferable kind of talk during peer collaboration (Barnes in Mercer & Hodgkinson, 2008; see also Mercer & Howe, 2012). Although Barnes rather describes the context and characteristics of exploratory talk and does not really provide a closed definition, his 1976 and 2008 descriptions are still being quoted or referred to (e.g. Brevig, 2006; Brown, 2016; Bullen et al., 2002; Enghag et al., 2007; Hewitt, 2014; Lofgren et al., 2013; Soter et al., 2008).

Cumulative, disputational and exploratory talk (the SLANT research project)

The onset to a more precise definition of exploratory talk was given during the two-year SLANT research project which investigated the potential of computers as a medium for exploratory talk in the primary classroom (Fisher, 1993). Pupils' dialogues were analysed extensively and, as mentioned in section 3.1.1, it was found that their talk fell generally under three headings: *cumulative talk, disputational talk* and *exploratory talk*. According to Fisher disputational talk 'can be characterized as an initiation in various forms (e.g. suggestion, instruction), followed by a challenge [resulting] either in a lack of any clear resolution or a resolution which does not build directly on the previous utterance' (Fisher, 1993, p. 255). In cumulative talk 'initiations are accepted either without discussion or with additions or superficial amendments' (Fisher, 1993, p. 255). Both cumulative and disputational talk appear to have little potential for learning (Fisher, 1993). This is different for exploratory talk, 'in which the initiation may be challenged or counter-challenged, but with suggestions which are developments of that initiation. Progress then rests on the joint acceptance of one of the suggestions, or of a modification of what has been put forward' (Fisher, 1993, p. 255).

Towards a standard definition of exploratory talk

The findings of the British SLANT project inspired other researchers to further explore exploratory talk for learning, both theoretically and experimentally. Since then the triad cumulative-disputation-exploratory has been referred to (as a taxonomy) in almost every research on pupil-pupil talk in the classroom. Building on the findings of the SLANT projects and on Barnes (1976), Mercer (1995) defined exploratory talk as

> the kind of talk in which 'partners engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counterchallenged, but challenges are justified and alternative hypotheses are offered. […] Knowledge is made publicly accountable and reasoning is more visible in the talk. Progress then emerges from the eventual joint agreement reached' (Mercer, 1995, p. 369; also cited in Wegerif & Mercer, 1997a, p. 53).

Wegerif and Mercer (1997a) also described exploratory talk as a 'style of interaction which combines reasoning through talk involving identifiable hypotheses, challenges and justifications, with a cooperative framework of ground rules emphasising the shared nature of the activity and the importance of the active participation of all involved' (Wegerif & Mercer, 1997a, p. 52). Within the same group of researchers Dawes, Dore, Loxley and Nicholls (2010) write: 'Exploratory talk happens when everyone is invited to give their ideas, and when children know how to challenge one another respectfully, sharing information, and giving and asking for reasons. Active listening is a major feature of productive group talk' (Dawes et al., 2010, p. 101).

Overall, Mercer's 1995 definition of exploratory talk is most being referred to. To illustrate this: his definition is quoted or referred to in all 63 articles reviewed. Barnes' 1976 pioneer work is referred to 41 times, while his description of exploratory talk is quoted or referred to 17 times. Barnes' 2008 description is quoted or referred to seven times.

In 2009, Enghag et al. formulated a more descriptive definition though not fundamentally different from Mercer's: 'We defined talk as exploratory if students have subject-matter focused talk and use language in an exploratory fashion, such as questioning, challenging, and encouraging. They often use half sentences and interject to fill missing words into the other person's sentences' (Enghag et al., 2009, p. 457). The same holds for Coultas (2012): 'This exploratory talk is the type of talk that leads to the guided construction of knowledge that can develop pupils' thinking. It involves sharing ideas and giving reasons for them and this is the kind of talk that allows for cognitive challenge and development' (Coultas, 2012, p. 176).

Atwood et al. (2010) elaborate on Mercer's definition by using more operational terms and describing some of the processes Mercer includes in his definition, though not adding any substantially new characteristics: 'Exploratory talk is that the methods used to reason are explicit; that is, these methods are observable in what participants do and are thus publicly accountable. Such methods include questioning of one's own and others' assumptions, outlining reasons for claims, making explicit evaluations and critiques, and engaging in persuasion. When challenges occur, participants give reasons and offer alternatives. Furthermore, challenges are launched from a stance in which the aim is to lay bare reasoning processes in order to make them available to others for the purposes of refining and reconstruction. As is evident in this description, exploratory talk is cooperative interaction' (Atwood et al., 2010, p. 366). And some researchers restrict themselves to the very basics of the definition, e.g. exploratory talk '[…] is talk that teachers and learners use when committed to learning and building understanding together' (Rutter, Edwards & Dean, 2016, p. 23).

In search of a hierarchical structure of types of pupil-pupil talk

Though implicitly confirming the added value of exploratory talk vs. cumulative and disputational talk, most researchers do not put this triad into a hierarchical structure. Exceptionally, and taking collaborative learning as a criterion for assessment, Schmitz and

Winskel (2008) consider disputational talk to be the lowest and least valuable level of talk, because its orientation is basically competitive. Competition remains absent in cumulative talk, but here the orientation is solitary. Therefore, Schmitz and Winskel argue, cumulative talk is the second level of talk. Exploratory talk is the highest level as its orientation is 'working towards best solutions through shared reasoning' (Schmitz & Winskel, 2008, p. 583).

Finally, some researchers use the term exploratory talk as a criterion to analyse pupil-pupil talk but they do not specify what definition of the term is being used. Mostly this is because they use different frameworks or typologies in which the notion 'exploratory talk' is primarily used to describe an aspect or dimension of that framework. Murcia and Sheffield (2010) do not refer to Barnes or Mercer but categorise student talk within the Mortimer and Scott (2003) framework 'Dimensions of discourse and the communicative approaches'. In their analysis of the quality of student talk, Murcia and Sheffield distinguish between 'Argumentation-reasoning', 'Exploratory talk', 'Student to student' and 'Other (low quality)'. It is unclear to us why a distinction is made between argumentation-reasoning and exploratory talk, as according to Barnes (1976) and Mercer (1996) the first is by definition included in the latter.

Summarising, researchers seem to agree that exploratory talk is about peer collaboration/interaction, visible reasoning, knowledge construction and public accountability. This is realised by means of active and respectful listening, critically challenging or counter-challenging ideas, tentatively discussing hypotheses, giving and asking for arguments and striving for joint agreement. Ground rules function as an underlying framework for this type of talk.

In order to further define exploratory talk we will look at (near) synonyms, antonyms, refining, elaborating and overarching terms. In addition we will summarise how the concept of exploratory talk has evolved.

3.1.2.2 Synonyms and antonyms

Our database search has provided a considerable number of (near) synonyms of exploratory talk. We also found a few antonyms. Table 5 summarises our findings.

Table 5
(Near) synonyms and antonyms for exploratory talk[5]

Synonyms	**Source**	**Antonyms**	**Source**
accountable talk	Michaels, O'Connor & Resnick, 2008; Webb, Whitlow & Venter 2016; Gillies 2014	adversarial discourse	Nussbaum, 2005
collaborative reasoning	Webb et al., 2016	expository talk	Holmes, 1992; Yaguchi, 2008
critical discussion	Webb et al. 2016		
dialogic argumentation	Gillies 2014		
dialogic discussion	Gillies 2014		
dialogic talk	Wegerif 2013; Chick 2015; Murcia & Sheffield 2010; Lofgren et al., 2013; Boyd & Kong, 2015		
exploratory discourse	Brown, 2016; Fernandez et al., 2001; Golanics & Nussbaum, 2008; Kumpulainen, 1996; Nussbaum, 2005		
interactive talk	Lofgren et al., 2013; Murcia & Sheffield, 2010		
reflective talk	Chick, 2015; Nikolaidou, 2012		
transactive reasoning	Berkowitz & Gibbs, 1983; Kruger, 1993		

In this section we will round up what the reviewed articles revealed. We do so in a chronological order to see if, in the course of time, the concept has become subject to either convergence or divergence.

One synonym is *transactive reasoning*, a term introduced by Berkowitz and Gibbs (1983), inspired by Bentley and Dewey (1949). Berkowitz and Gibbs define *transactive reasoning* as 'reasoning that operates on the reasoning of another. […] In a very dialectical sense, one's own reasoning confronts the other's antithetical reasoning in an ongoing dialogic dynamic' (Berkowitz & Gibbs, 1983, p. 402). Kruger (1993) also uses this term and makes its definition more operational by displaying concrete elements: '[…] criticisms, explanations, justifications,

[5] Some authors refer to previously published sources which are not the subject of this study. They are therefore not mentioned in this overview.

clarifications, and elaborations of ideas.' In her experiment she found that 'peers directed these exchanges at each other; they transacted on each other's ideas' (Kruger, 1993, p. 167). Though close to exploratory talk as far as speech acts are concerned, the term transactive reasoning has not been used again in the articles analysed for this review, except for short references by Wegerif and Mercer (1997a) and Schmitz and Winskel (2008). Perhaps the fact that speaking of transactive reasoning requires a more dialectic view on peer group conversation is considered too determining, as not all pupils enjoy a dialectic approach to learning (Robins, 2011) and exploratory talk is more than dialectic only (Wegerif, 2013).

Chick (2015), referring to Mercer (2000), uses the notion of *dialogic* or *reflective talk*, defining it as 'talk which is characterized by features such as constructive engagement with each other's ideas, a spirit of enquiry and intellectual openness, and by an atmosphere of trust. It is a type of talk where suggestions can be offered for joint consideration and opinions treated with respect' (Chick, 2015, p. 299). The notion of reflective talk was only picked up by Nikolaidou (2012), who distinguishes it from exploratory talk (cf. infra). Mannion and Mercer (2016) introduce exploratory talk in a 'reflective' learn-to-learn project, during which pupils were required to develop their ability to reason out loud, thinking and working together in pairs and small groups. It did not make both researchers replace the term exploratory talk with *reflective talk*, though, which makes us presume that, to avoid confusion, Mercer prefers to use the term 'exploratory talk'.

Some researchers occasionally use the term *exploratory discourse* as a synonym of exploratory talk (Fernandez et al., 2001; Kumpulainen, 1996; Nussbaum, 2005), but only Golanics and Nussbaum (2008) do this consistently. While they describe exploratory discourse as 'the functional equivalent of collaborative argumentation', they seem to concur with Mercer's definition of exploratory talk: 'Our definition of collaborative argumentation is similar to Mercer's (1996) notion of exploratory discourse' (Golanics & Nussbaum, 2008, p. 168). In collaborative argumentation, students work together to construct and critique arguments (Golanics & Nussbaum, 2008). Collaborative argumentation, Wegerif, Mercer and Dawes (1999) confirm, promotes more complex and critical thinking. By using the word *discourse*, Golanics and Nussbaum fortify the argument that exploratory talk is a characteristic of interaction at group level. The term *exploratory discourse* is also used once, though undefined, by Fernandez et al. (2001).

Like Murcia and Sheffield (2010), Lofgren et al. (2013) refer to the analytical tool developed by Mortimer and Scott (2003) in order to evaluate communication and meaning-making processes in the classroom. The tool focuses on the scaffolding efforts of the teacher, i.e. the way in which he supports pupils in developing their knowledge. It is characterised by four classes, one of which is called *interactive/dialogic talk*. About the latter, Lofgren et al. (2013) write: 'There are strong similarities with Barnes' (2010) definition of exploratory talk, which incorporates different ideas or opinions, explicit reasoning, critical but constructive engagement' (Lofgren et al., 2013, p. 486).

According to Boyd and Kong (2017), Mercer and his colleagues now call exploratory talk *dialogic talk*. To make this claim they refer to Wegerif (2013), who found inspiration for the idea in an experimental study in Mexico. There, Rojas-Drummond and Zapata (2004) had examined whether the teaching of exploratory talk without any explicit reasoning can also lead to an improvement in collaborative, creative or divergent tasks. And so, Wegerif argues: 'What is essential to exploratory talk is not in fact the explicit reasoning [as invoked by the definitions of Barnes and Mercer …]. Just as disputational talk and cumulative talk can best be defined by the type of identification they imply, so can the intersubjective reality referred to previously by the term exploratory talk. I now prefer the term dialogic talk since what seems to be most essential to this type of talk is identification with dialogue itself' (Wegerif, 2013, p. 51).

Zooming in on the aspect of identity, Polo, Lund, Plantin and Niccolai (2015) acknowledge that disputational, cumulative and exploratory talk reflect different attitudes towards self-identity at the individual level. Nevertheless, as post-2013 research is still using the term exploratory talk abundantly, it would be too soon to just do away with it. Moreover, dialogic talk is not a new concept. It was introduced by Vygotsky's contemporary theoretician Bakthin who claimed that language is a social practice and all thought is dialogic (Bakhtin, 2010; see also Lyle, 2008). The concept of dialogic teaching was rejuvenated by Alexander (2001), who describes it as what happens when teachers and pupils work together to build on their own and each others' knowledge and ideas in order to develop coherent thinking. For Alexander (2008), dialogic teaching reflects a view that knowledge and understanding come from testing evidence, analysing ideas and exploring values, rather than accepting somebody else's certainties without questioning them. Alexander's (2008) definition of dialogic talk is that it should be collective, reciprocal, supportive, cumulative and purposeful. Although there are many similarities with exploratory talk, we find it premature to call it a synonym of dialogic talk, because, as Sutherland (2013) puts it: 'Exploratory talk has many features in common with 'dialogic' talk, but it also stresses public accountability and the visibility of reasoning processes' (Sutherland, 2006, p. 45). As if to emphasise the difference, Alexander (2010) has added exploratory talk, together with expressive talk, as a fifth and sixth kind of talk which can help teachers to engage in dialogic teaching.

Finally, according to Webb et al. (2016), exploratory talk 'is not far removed from *collaborative reasoning* (Chinn & Anderson 1998; Reznitskaya et al. 2009), *critical discussion* (Keefer et al., 2000), *accountable talk* (Michaels & O'Connor, 2008), and argumentation approaches to learning science (Osborne, 2010)' (Webb et al., 2016, p. 4). It shares conceptual and procedural features with dialogic argumentation and dialogic discussion, etc. Gillies (2014), too, refers to the concept of *accountable talk.* Michaels et al. (2008) describe it as talk that emphasises logical connections and the drawing of reasonable conclusions. It involves explanation and self-correction. It often involves searching for premises, rather than simply supporting or attacking conclusions. And 'speakers make an effort to get their facts right and make explicit the evidence behind their claims or explanations. They challenge each other when evidence is lacking or unavailable' (Michaels et al., 2008). Gillies (2014), following Alexander (2010), also sees parallels between accountable talk and dialogic teaching.

As far as antonyms are concerned, Nussbaum (2005) opposes exploratory discourse to *adversarial discourse* and refers to Wegerif et al.'s findings that it is the kind of talk most closely linked to learning outcomes (Wegerif, Mercer & Dawes, 1999). Further, Brown (2016) seems to use the term as a synonym when she writes: 'The dialogue then lends itself to introducing a counter-argument, and continues with exploratory discourse of 'I believe' or 'I think' using generalisations to create and open up dialogue' (Brown, 2016, p. 88).

Yaguchi, Iyeiri and Baba (2010) adopt the term *expository* vs. *exploratory* from Holmes (1992) who explains that expository talk conveys facts and/or opinions while exploratory talk develops ideas through negotiation. In an educational context this comes close to Barnes' classroom talk dichotomy of *presentational* vs. *exploratory* talk (cf. supra). Yaguchi adds that in exploratory talk 'the speaker shows an affective attitude toward the listener' (Yaguchi et al., 2010, p. 587).

We discussed (near) synonyms and antonyms each in chronological order to find any signs of convergence or divergence. It appears that most (near) synonyms and antonyms as well are fairly recent, which suggests divergence in (recent) time. This seems paradoxical as the definition or description of exploratory talk has not changed that much. We therefore believe that (near) synonyms and antonyms of exploratory talk mostly emerge from theoretical preferences or certain viewpoints on communication and education.

3.1.2.3 Elaborating/refining terms

Apparently, it has taken researchers some time to start deepening the concept of exploratory talk, for most of the elaborations we found are fairly recent. Table 6 provides an overview.

Table 6
Refining and elaborating terms for exploratory talk[6]

Elaborating/refining terms	**Source**
group exploratory talk	Sutherland, 2006
incipient vs. elaborate exploratory talk	Rojas-Drummond et al., 2003
inclusive vs. exclusive exploratory talk	Rajala, Hilppo & Lipponen, 2012
reflective and operational talk	Brevig, 2006; Nikolaidou, 2012
replacement, interweaving, contextual privileging and pastiche	Renshaw & Brown, 2007

[6] Some authors refer to previously published sources which are not the subject of this study. They are therefore not mentioned in this overview.

In a replicator study Rojas-Drummond et al. (2003) distinguish between *incipient exploratory* and *elaborate exploratory talk*. Incipient 'suggests exploratory talk is neither very consistent nor very prominent in the way children talk, whereas [elaborate] indicates exploratory talk is more consolidated and sophisticated' (Rojas-Drummond et al., 2003, p. 359). This distinction has implicitly been acknowledged by Herrlitz-Biro, Elbers and de Haan (2013), who argue that the analysis of key words as indicators of exploratory talk ('Why?', 'What do you think?', 'because', etc) ignores other aspects of exploratory talk, such as collaborative processes. In an experimental study they did a qualitative analysis which demonstrated that pupils can talk exploratively without using such 'key words'. They consider elaborations, i.e. the restructuring of information or linking new information to existing information to be a main ingredient of exploratory talk but do not go so far as to claim that exploratory talk is the same as *elaborating talk* (see also van Boxtel, van der Linden & Kanselaar, 2000). Nevertheless, parallels can be drawn with the lack of prominence and consistency of exploratory talk in the Rojas-Drummond et al. study. Thus far, though, the concepts of incipient and elaborate exploratory talk have not been re-used by other researchers.

Sutherland (2006) refers to *group exploratory talk*, which is 'characterized by equality of participation, with pupils responding to each other's points; 'tentativeness' (Wilkinson et al., 1965, cited in Howe, 1997); pupil higher-order questions or statements, requiring reflection or 'wait' time (Tobin, 1987); requests for clarification or illustration; and an ability to elaborate and sustain the dialogue (Barnes & Todd, 1977; Mercer, 2000)' (Sutherland, 2006, p. 107).

In an analytical framework for integrating every day and scientific discourse Renshaw and Brown (2007) identify four formats of classroom talk: *replacement, interweaving, contextual privileging* and *pastiche*. They consider these as 'different instantiations of exploratory talk and a useful differentiation of that style of classroom talk' (Renshaw & Brown, 2007, p. 543).

Introducing the notions of *inclusive* and *exclusive exploratory* talk Rajala et al. (2012) focus on quantitative and interactional (a)symmetry, i.e. the participation rate of individuals in group discussion. When all pupils are able to more or less equally contribute to exploratory discussions (quantitative symmetry) Rajala et al. speak of *inclusive* exploratory talk. When this is not the case, e.g. when one pupil drops out of the discussion or starts dominating the conversation, exploratory talk is *exclusive*. Referring to earlier studies Rajala et al. (2012) consider quantitative (a)symmetry an important aspect of collaborative learning and also of exploratory talk. They justify their claim by referring to Mercer's 'ground rules' for exploratory talk, which include that pupils encourage each other to be involved (see section 3.2 for a full discussion of ground rules).

In the context of his research Nikolaidou (2012) finds the triad cumulative-disputational-exploratory insufficient. The author adds the notions of *reflective* and *operational talk* to construct an extended matrix. In *operational talk* 'utterances relate to operational transactions with regard to talk and software respectively'. *Reflective talk* means 'engaging critically and constructively expressing a self-reflective thinking' (Nikolaidou, 2012, p. 744). The term operational talk does not reappear in the articles we reviewed. Reflective talk is mentioned by Chick (2015, cf. supra) and in an experiment during which students used exploratory talk to discuss literature, Brevig (2006) also makes a link with reflection: ‘I am confident that exploratory talk and reflection assist students in developing meaning. They can self-monitor their learning and develop and nurture evolving ideas’ (Brevig, 2006, p. 529).

3.1.2.4 Overarching terms

Some studies suggest broader concepts or terms that comprise exploratory talk. In our review study, we found several such concepts (see Table 7), which we will discuss in their chronological order of appearance.

Table 7
Overarching terms for exploratory talk[7]

Overarching terms	**Source**
collaborative argumentation	Golanics & Nussbaum, 2008; Herrlitz-Biro et al., 2013
co-constructive argumentation, critical argumentation, productive & accountable talk; collaborative argumentation	Polo et al., 2015
co-constructive talk	Rojas-Drummond, Mazon, Fernandez & Wegerif, 2006
collaborative engagement	Atwood et al., 2010
constructive talk	Dourneen, 2013
constructively-critical footing	Polo et al., 2015
critical learning	Riley, 2006
dialogic reasoning	Dourneen, 2013
IDRF – Initiation-Discussion-Response-Feedback	Wegerif & Mercer, 1997a
intellectual estuary	Beghetto & Kaufman, 2009
public discourse	Harris & Ratcliffe, 2005
quality talk	Atwood et al., 2010
shared thinking	Bowskill, 2010

IDRF, or Initiation Discussion Response Feedback, may be considered as one of the first overarching terms. It refers to Sinclair and Coulthard's (1977) I-R-F pattern, where I-R-F stands for Initiation – Response – Feedback. As discussed extensively in Chapter 2, section 2, it has been found as a dominant discourse pattern in the classroom by dozens of studies. Over the years it has also become clear that it does not make pupils learn through talk. Therefore, the

[7] Some authors refer to previously published sources which are not the subject of this study. They are therefore not mentioned in this overview.

interaction needed between teacher and pupils is not an I-R-F pattern but an IDRF-pattern (Wegerif and Mercer, 1997a), where D stands for 'Discussion'. IDRF ‘describes the basic structure of the educational activity of groups working together […]. Where the discussion element is exploratory, this exchange structure combines an aspect of directive teaching with an aspect of exploratory learning’ (Wegerif, 1996a, p. 13).

Harris and Ratcliffe (2005) integrated exploratory talk in a project to explore socio-scientific issues relating to genes and genetic engineering in secondary schools. They refer to three different models of teaching citizenship, as introduced by Huddleston & Rowe (2003): the ‘civics’ model (transmission of knowledge), the ‘current affairs’ model (more or less critical exchange of opinions) and the ‘*public discourse’ model*. In the latter pupils explore where their views come from and critically examine the views of others, so that discussion is not merely the activity but a means of developing the ability to participate in informed public discourse. To Harris and Ratcliffe (2005), public discourse implies the use of exploratory talk, though they also use both terms as synonyms.

Acknowledging Mercer’s definition of exploratory talk, Rojas-Drummond et al. (2006) suggest that we use one single overarching term, i.e. *co-constructive talk* (and also co-constructive interaction) ‘as an inclusive term to characterise the joint efforts of coordination, negotiation and collaboration in various group work activities’. They add that exploratory talk is ‘a particularly effective and sophisticated type of educated talk or social mode of thinking, which represents one specific form of co-constructive interaction’ (Rojas-Drummond et al., 2006). Their argument is that the contrast between exploratory and cumulative talk is too artificial, as it depends on how one defines ‘explicit reasoning’. This overarching concept would also be useful to ‘characterize a much wider scope of collaborative discussions children display when working together to solve problems […] and in many educational contexts’ (Rojas-Drummond et al., 2006, p. 93).

Riley (2006) sees exploratory talk as an exponent of *critical learning*, which he describes as ‘a reflective activity with critical intent that enables students to socially engage in learning tasks and collaborative problem solving through sharing and challenging personal perspectives, experiences and knowledge to co-construct knowledge and generate solutions and outcomes by using peer-critical evaluation and reflective practices’ (Riley, 2006, p. 63).

Enghag et al. (2009) consider group discussions and exploratory talk as indicators of *group ownership*. Within this context Haglund and Jeppsson (2012) characterise exploratory talk as ‘rapid turn-taking, incomplete sentences and sustained focus on a shared line of reasoning’ (Haglund & Jeppsson, 2012, p. 910), all characteristics previously made visible by Barnes (1976) and Mercer (1995).

Beghetto and Kaufman (2009) plunge exploratory talk in an *intellectual estuary* which ‘describes an area of great and diverse intellectual identities in which separate streams of ideas flow in and meet with the vastness of ideas found in a given academic discipline’ (Beghetto &

Kaufman, 2009, p. 312). In fact, they argue that several characteristics of exploratory talk are indicative for the formation of an intellectual estuary: the use of ground rules (make claims using arguments, challenge claims with counter-arguments, strive for agreement), orientation to shared reasoning and using opportunities to create and support new insights.

Somewhat similar to Beghetto and Kaufman's intellectual estuary, Bowskill (2010) has developed a socio-cultural practice and theoretical framework for 'a new generative learning environment that creates shareable electronic artefacts from reflective dialogue across a whole-group' (Bowskill, 2010, p. 61). He calls it *shared thinking*. 'This environment contains a space, a structure, a reflective dialogue, a disposition, a purpose and a shareable product' (Bowskill, 2010, p. 61). The dialogical space (see also Mercer, 2000), in which students generate question options by reflecting upon and sharing their experiences, is generated by a protocol resembling exploratory talk. To Bowskill (2010), a 'listening pedagogy' is one of the important characteristics of shared thinking.

Atwood et al. (2010) use Crook's (1998) notion of *collaborative engagement*, which is comparable to *collaborative argumentation* (Golanics & Nussbaum, 2008; Herrlitz-Biro et al., 2013), or *exploratory discourse* (cf. supra). Basically, it is synonymous to Beghetto and Kaufman's *intellectual estuary* and there are also some parallels with Riley's *critical learning*. As to Atwood et al., collaborative engagement describes a classroom in which students engage one another's ideas through joint or collective reasoning. Exploratory talk, then, is the specific form of talk learners use to co-construct their reasoning process (Brown & Renshaw, 2000). Atwood et al. (2010) also refer to the term *quality talk*: 'Mercer (1995, 2000), Mortimer and Scott (2003), and Van Boxtel and Roelofs (2001) have characterised quality talk as that which displays reasoning, the articulation of propositions, and the clarification of misconceptions about those propositions' (Atwood et al., 2010, p. 363).

Dourneen (2013), too, is not convinced the term exploratory talk is adequate enough. In a qualitative study focusing on pragmatic features while students perform tasks, she argues that exploratory talk might not always explain how learners make sense of each other in order to develop their ideas. She would rather speak of *dialogic reasoning*, quoting Wegerif et al. (2005), as 'the broader concept of shared orientation, ground rules and utterances that helps people reach shared understandings and construct shared new knowledge' (Wegerif et al., 2005, p. 43 in Dourneen, 2013). In addition *constructive talk* is suggested as a concept, meaning 'talk which enables learners to construct ideas and helpful working relationships through the pragmatic expressions they use to make meaning in the context in which they are working' (Dourneen, 2013, p. 46).

Finally, Polo et al. (2015) bring together four specific types of talk which are considered to be of high educational value: *academically productive talk, accountable talk, collaborative argumentation* and *co-constructive, critical argumentation*. One may argue to what extent these concepts are in fact over-arching, but according to Polo et al. (2015) exploratory talk shows characteristics of each of these types. Zooming in on the aspect of identity, he acknowledges Mercer's (1996) statement that in exploratory talk there is no conflict of people, only conflict of ideas. According to Polo et al. (2015), disputational, cumulative and exploratory talk reflect different attitudes towards self-identity at the individual level. In the case of disputational talk Polo et al. (2015) speak of 'competitive footing' and in the case of cumulative talk it is 'consensual footing'. For exploratory talk they speak of 'constructively-critical footing', where 'footing' 'corresponds to the changing roles an individual displays during a conversation' (Polo et al., 2015, p. 6).

3.1.3 Contextual use

The context in which exploratory talk has added value has been suggested by a number of studies in terms of psychological, social, cognitive and educational significance. Though the borderlines between those dimensions are not always clear, we will attempt to discuss them separately.

Psychological

An extensive contribution to the psychological dimension of exploratory talk is made by Polo et al. (2015). They relate the three types of talk which Dawes, Fisher and Mercer (1992) named - disputational, cumulative, exploratory - to three types of recognition students may expect, corresponding to what they base their self-identity on. In disputational and cumulative talk, self-identity is challenged. Cumulative talk creates solidarity with the group, while disputational talk makes participants create their self-identity at the cost of the self-identity of others. In exploratory talk, however, self-identity becomes irrelevant, as it is transferred to the group level. Nobody in the group has to lose his face. In this respect, exploratory talk offers the possibility to create shared ownership or, as Enghag et al. (2009) call it, 'group ownership', which 'refers to the groups' choice and control of the management of the task and how the task is determined, performed, and finally reported' (Enghag et al., 2009, p. 456-457).

In a study involving a reading project, Mayher (1990) highlights the reflective power of exploratory talk. Quoted in Brevig (2006, p. 523), he suggests: 'Exploratory talk allows us to check our emerging sense of what we've read or heard or seen by sharing our perceptions with others and learning further from their similar explorations'. Wegerif et al. (2005) link exploratory talk to change of personal attitude, as its presence and mastering enable pupils to change their minds in response to good arguments. Nikolaidou (2012) suggests that the use of exploratory talk may benefit the individual's creative thinking.

Social

In the Vygotskyan tradition and building on Habermas' 'communicative rationality', Mercer (1995) makes clear that language is used by the group as a social mode of thinking. 'Exploratory talk typifies language which embodies certain principles of clarity, of constructive criticism and receptiveness to well-argued proposals in many societies' (Mercer, 1996, p. 370), making it especially relevant in contemporary domains as science, law, government and the negotiation of business (Herrlitz-Biro et al., 2013; Mercer, 1996; Mercer et al., 1999; Wegerif & Mercer, 1997a). Herrlitz-Biro et al. (2013) add that through exploratory talk pupils can develop 'explicit multi-perspectivity', which implies a more open-minded attitude. When confronted with different views such an attitude has a lot of potential to create and establish new insights. This potential was also acknowledged by Luby (2014) who used interreligious dialogue to promote the development of both cumulative talk and exploratory talk amongst secondary school pupils.

Cognitive

Suggesting it as an equivalent of collaborative argumentation, Golanics and Nussbaum (2008), state that exploratory talk 'has been found to deepen subject matter understanding (Alexopolou & Driver, 1996; Bell & Linn, 2000) and cause conceptual change (Schwarz et al. 2000; Baker, 2003). In addition to knowledge building, collaborative argumentation promotes more complex and critical thinking (Wegerif et al., 1999) [when] critical thinking can be defined as the ability to identify, construct and evaluate arguments (Finocchiaro, 2005)' (Golanics & Nussbaum, 2008, p. 168).

Educational

Exploratory talk is of central concern and even most productive in learning situations, many researchers agree (Barnes, 1976; Brevig, 2006; Gillies, 2014; Mercer, 1996; Mercer et al., 1999; Rojas-Drummond & Zapata, 2004; Tin, 2003; Topping & Trickey, 2014; Webb, 2015; Wegerif, Mercer & Dawes, 1999; Wheeldon, 2006). As it increases reasoning skills (Golanics & Nussbaum, 2008; Mercer et al., 1999; Robins, 2011; Rojas-Drummond et al., 2003; Webb & Treagust, 2006; Webb et al., 2016; Wegerif & Mercer, 1997a), it maximises the potential for learners to construct shared meanings and reach agreements, allowing for collective problem solving. Mercer (1996) finds exploratory talk most effective for solving problems through collaborative activity.

Based on a qualitative analysis of conversations of 10- and 11-year-old pupils studying global citizenship through an online discussion environment, Riley (2006) found exploratory talk to be an indicator of critical learning, as the latter requires collaborative learning approaches and the application of key skills in order to develop reflective activities. Nikolaidou (2012) considers exploratory talk as a means to create transfer between subjects as pupils become better to develop intellectual habits that will furnish them well across various learning situations. In this sense, exploratory talk is most productive in learning situations (in general and at school), because it increases reasoning skills.

Finally, the educational benefit also serves the teachers, as they are encouraged to teach more dialogically and to create an interactive environment (Webb, 2015).

3.1.4 Conclusion

Stimulated by a revival of Vygotskyan learning theory, research on the connection between language and learning has strongly increased over the last four decades. There has been and still is a strong agreement about the value of collaborative learning in education and of the role of particular kinds of talk in the process. This literature review focused on the notion of exploratory talk, which was launched by Barnes in 1976 and - as part of a taxonomy which also includes disputational and cumulative talk - defined by Mercer nineteen years later.

Our first research question is: how is exploratory talk defined? Subquestions are: in what way has the concept of exploratory talk evolved in terms of its definition and characteristics (concept growth)? Have any synonyms been introduced? Have any elaborating, deepening and/or overarching terms been proposed? Which definition of exploratory talk is currently being used? In order to answer these questions we thoroughly analysed 63 articles selected from 6 databases.

Barnes (1976, 2010) launched the notion of exploratory talk as a means for learning but never formulated a real definition. Building on Barnes' work Mercer (1996) formulated a definition to which only few characteristics have been added or perhaps it is more correct to say some have been 'highlighted'. Recapitulating, Mercer defines exploratory talk as the kind of talk in which 'partners engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counterchallenged, but challenges are justified and alternative hypotheses are offered. […] Knowledge is made publicly accountable and reasoning is more visible in the talk. Progress then emerges from the eventual joint agreement reached' (Mercer, 1996, p. 369; Wegerif & Mercer, 1997a, p. 53). This definition has been quoted or referred to in all articles reviewed. Elaborating on it, Yaguchi et al. (2010) stress the affective attitude toward the listener, an aspect Dawes et al. (2010) emphasise by calling active listening a major feature of productive (exploratory) group talk. In turn, Enghag et al. (2009) stress the aspect of 'subject-matter focus' as a main characteristic. Both listening and subject-matter focus seem to us operational clarifications of Mercer's definition rather than substantial additions.

As we have seen, there have been some elaborations and refinings, and synonyms for the concept have been proposed, but the definition itself has not been refuted. This makes us conclude that it is still valid and is considered, more than Barnes' descriptions, to contain a workable set of criteria for analysing classroom talk.

Though synonyms are rather scarce we believe the following notions share nearly the same characteristics: *exploratory discourse, collaborative argumentation/reasoning* and maybe also *dialogic talk* and *transactive reasoning*, although in the eyes of Sutherland (2013) both would lack the aspects of public accountability and the visibility of reasoning processes. Very close to exploratory talk, then, come the notions of *accountable talk* and *critical discussion*.

Elaborating or refining terms of exploratory talk have been suggested: *incipient* and *elaborate exploratory talk* were introduced by Rojas-Drummond and Mercer (2003) and implicitly acknowledged by Herrlitz-Biro et al. (2013). Sutherland (2006) talks of *group exploratory talk*, focusing on the equality of participation, a notion further explored by Rajala et al. (2012) who distinguish between *inclusive* and *exclusive exploratory talk*. Nikolaidu (2012) adds *reflective* and *operational talk* to the triad cumulative-disputation-exploratory talk. Hardly none of these suggestions have been followed, so far.

As far as overarching terms are concerned, discussion has been rather limited though many suggestions have been made, e.g. *IDRF* (Wegerif & Mercer, 1997a), *co-constructive talk* (Rojas-Drummond et al., 2006), *critical learning* (Riley, 2006), *collaborative argumentation* (Golanics & Nussbaum, 2008), *dialogic reas*on (Dourneen, 2013). Polo et al. (2015) mention four terms of which exploratory talk shares characteristics, but no more than that. It is our impression that researchers do not really seem to agree or disagree with one another and neither does there seem to be a common ground for the one or the other. Actually, researchers rather seem to 'resolve to cumulative talk' about the topic. Suggestions seem to be based on theoretical preference, not on theoretical need.

Essentially, in an educational context, the concept of exploratory talk has found solid ground in two stages: Barnes (1976) and Mercer (1996). We conclude that the notion of exploratory talk has evolved from a mere type of talk to an indispensable part of a triad taxonomy of classroom talk: disputational – cumulative – exploratory. Although Mercer's definition of exploratory talk still stands, the results of some studies suggest further refining or rather emphasising certain characteristics which we can agree with. We believe turn-taking and especially the lack of interactive dominance/recession to be important. We also believe group identity is a condition as well as a result of exploratory talk. Therefore we would like to elaborate on Mercer's definition as follows (the additions are marked in italics):

> Exploratory talk is *a specific form of co-constructive interaction expressed through discourse*, in which partners *equally participate to maintain a sustained focus on a shared line of reasoning*. They engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counterchallenged, but challenges are justified, and alternative hypotheses are offered. Knowledge is made publicly accountable. *Reasoning is more elaborate* and more visible in the talk. Progress then emerges from the eventual joint agreement reached *and supported by group ownership*.

3.1.5 Discussion

In this final section, we will sum up four issues that emerged from the reviewed articles as far as the definition, the concept growth and the contextual use of exploratory talk is concerned.

We will elaborate on these issues in the final chapter of this dissertation, i.e. in the section About defining exploratory talk.

- Each type of talk has its own value, but this value is context dependent and – in an educational context – task/target dependent.
- ‘Explicit reasoning’ must further be defined in order to preserve the taxonomy cumulative – disputational – exploratory talk.
- Before adding other kinds of talk to the triad taxonomy it must be clear which criteria have to be met in order for them to ‘fit in’.
- From a pragmatical point of view, the notion of exploratory talk seems more transparent to work with than the notion of dialogic talk.

3.2 The measurement of exploratory talk (RQ 1.2)

In this section we will discuss how exploratory talk is measured. We will present an overview of the type of studies we found, which methods were used and which variables and characteristics concerning exploratory talk were examined. However, before doing so, we believe it is useful to start with a more general methodological question: how is classroom talk measured and evaluated in general?

3.2.1 Linguistic ethnography and sociocultural research

In a review on methods and methodology Mercer (2010a) discusses relevant methods for analysing classroom talk and compares their strengths and weaknesses. Mercer distinguishes between two important approaches: linguistic ethnography and sociocultural research. Linguistic ethnography finds its rationale in social anthropology and descriptive linguistics. Researchers emphasise, among other aspects, ‘that talk is always referential, interpersonal, emotive and evaluative [and] that socialization is a never-ending process, mediated through language and interaction’ (Mercer, 2010, p. 2). As social situations are thought to be unique, ethnographic researchers refrain from quantitative approaches as these are often used to draw generalising conclusions which they do not believe to be valid. Therefore, within linguistic ethnography, studies are observational, non-interventional and qualitative. Researchers examine classroom talk in its social and cultural context through detailed close reading of transcripts and leave out statistical analysis. Research questions address issues such as the expression of identities by means of classroom talk and the use of languages and language varieties of different cultures at school.

Sociocultural researchers bear on research traditions in social and developmental psychology and pedagogical studies. They strongly attach to the ideas of Vygotsky (1978) who considers language to be a cultural and psychological tool. To Vygotsky and - in extenso - to sociocultural researchers knowledge and understanding are created together. Language is used to learn by joint reasoning (intermental) and what is learned is eventually integrated into the individual mind (intramental). Since education is considered to be a dialogic process, the way talk – i.e. dialogue – is organised in the classroom could have an important influence on pupils' reasoning and reasoning skills. Therefore sociocultural researchers 'are positively inclined towards the use of pre/post interventional designs, seeking to measure differential effects of talk on problem solving, learning and conceptual change' (Mercer, 2010a, p. 3). While often combining qualitative and quantitative methods, their studies are mostly observational, interventional and/or quasi-experimental. Research questions focus on the occurrence of types of classroom talk and the way these types promote learning and develop understanding.

Quantitative, qualitative and mixed methods to study classroom talk

Apart from both approaches, i.e. linguistic ethnographic and sociocultural, a distinction must be made between quantitative and qualitative methods for analysing talk. Briefly speaking, qualitative methods aim to 'reveal the nature, patterns and quality of spoken interactions' (Mercer, 2010a, p. 6). Data are approached bottom-up. In the linguistic ethnographic approach this means that classroom talk is analysed in depth and the amount of data is therefore rather restricted: a limited number of talk is audio and/or video recorded, transcribed and illustrative extracts from these transcriptions are analysed and their characteristics described. While being a very good method to examine talk in detail, qualitative methods are hard to use when dealing with a large amount of data. Moreover, they do not offer a systematic basis to abstract and generalise from small extracts of transcript (Wegerif & Mercer, 1997a).

Quantitative methods may include a set of characteristics of indicators being put into a coding scheme based on which talk features are classified, counted and the results interpreted statistically. This method is traditionally called 'systematic observation' (Croll, 1986; Teasley, 1995; Mercer et al., 2004). For indicators to be valid, transcripts may first be coded independently by two or more researchers, generating a score for interrater reliability and - in many cases - resulting in a level of agreement. This is necessary when e.g. counting certain words as an indicator for a type of talk alone is inadequate because contextual usage of those words may differ. Current statistical analyses in quantitative research of talk are frequency counts, significance and correlation techniques, and multilevel analyses. The disadvantage of a merely quantitative approach in the analysis of talk, however, is that the context of talk is difficult, if not impossible to incorporate.

Finally, Mercer (2010a) mentions conversation analysis (CA; see Schegloff and Sacks, 1973, 1992) as a methodology rather than a method. It uses 'a very detailed and laborious style of analysis and sets very strict criteria for the kinds of interpretations which an analyst can make from the data of recorded talk, and it also involves the use of a very specific and detailed method of transcription' (Mercer, 2010a, p. 8). Coding methods enable researchers to correlate features of observed events with outcomes. However, 'they are weak in explaining any process by which such correlations may arise' (Wegerif & Mercer, 1997b, p. 276).

In a sociocultural approach qualitative methods of data collection and analysis may be applied as described above, but methods of descriptive linguistics may be added, e.g. analysis of prosodic elements, grammatical constructions, vocabulary items, etc. Quantitative methods may be used as well, a reason for which a larger amount of data is often collected and analysed semi-automatically, using specialised software like NVivo (Bazeley, 2013). Combining qualitative and quantitative methods results in 'mixed methods' research of the kind that is often used for studies in the fields of social psychology, sociolinguistics and education.

According to Mercer (2010a) only one mixed methods approach has actually been given a specific name: *sociocultural discourse analysis*. Where linguistic discourse analysis is much concerned with the organisational structure of spoken language, sociocultural discourse analysis focuses on language 'content, function and the ways shared understanding is developed, in a social context, over time' (Mercer, 2010a, p. 9). Furthermore, sociocultural discourse analysis differs from discourse analysis as the latter is used to refer to several different ways of analysing language (Mercer, Littleton & Wegerif, 2004). As such, sociocultural discourse analysis is very useful to analyse and evaluate the talk of pupils while working together in pairs or small groups. The qualitative aspect comes down to interpreting certain fragments of transcripts and integrating this analysis with a quantitative approach. For example, in order to establish whether pupils think in a critical way about what is being said during collaborative activities, a qualitative analysis may reveal the why-question as an indicator. Through quantitative analysis with concordancer software, such as NVivo, the frequency of this question in conversations can then be determined, e.g. before and after an intervention. Wegerif and Mercer (1997b) offer additional reasons to employ a mixed methods approach such as sociocultural discourse analysis, an important argument being that it is easier to replicate and compare such studies than e.g. studies in which only a pure qualitative approach is employed. They quote Hammersley (1992) saying: 'Qualitative analysis can be effective for generating theories but not so effective for rigorously testing them' (Wegerif & Mercer, 1997b, p. 275).

3.2.2 Quantitative and qualitative analysis

Knowing the chief methods and methodologies to analyse classroom talk, we may now address the question how exploratory talk is measured. In order to answer this question we did a close reading of the 115 articles we found in the databases. 88 of those articles discuss empirical research on exploratory talk, 27 are either theoretical (18) or methodological (7) contributions, or compare data from similar studies (2). In this paragraph we will give a brief overview of the methods that were described in the 88 experimental research articles. In the next paragraph we will discuss in detail those methods which we consider relevant for our own research, adding input from the seven methodological articles. Table 8 gives an overview of the methods we found.

Table 8

Methods used to analyse exploratory talk (88 articles)

	Qualitative	Quantitative	Mixed methods
Number of studies	41	2	45

On the whole, we found 41 studies to be qualitative, two to be quantitative and 45 to be mixed methods. 33 of the 41 qualitative analyses include solely conventional analysis (CA). This means that exemplary extracts of transcripts are analysed in depth (cf. infra). In eight studies talk is analysed via other protocols (e.g. Ten Have, 1999) or 'conventional discourse analysis' without reference to any specific protocol. In one of the quantitative studies and nearly all mixed methods studies conventional discourse analysis is used.

Analytical frameworks

Table 9 gives an overview of the frameworks or definitions researchers used to analyse exploratory talk.

Table 9

Analytical and theoretical frameworks, definitions and characteristics used to analyse exploratory talk (88 articles)

	Definition ET Barnes	Definition ET Mercer	CDE coding scheme	Specific ET characteristics	Other analytical framework
Qualitative studies	9	2	9	13	10
Quantitative studies	0	0	0	1	1
Mixed methods studies	2	1	23	12	14
Total	11	3	32	26	25

Theoretical frameworks or references for these analyses are diverse: Barnes' definition of exploratory talk is exclusively used in 11 studies, Mercer's in three. Additionally, Mercer's coding scheme for social modes of thinking (CDE, cumulative - disputational - exploratory) is used in 32 studies. It zooms mostly in on so-called 'key words in context', the length of utterances and other linguistic characteristics. We will discuss this further in the next paragraph. Some characteristics of exploratory talk are analysed separately in 26 studies. In 25 studies other frameworks related to exploratory talk are used. We will now look at these numbers more closely.

Qualitative studies

Theoretical frameworks or references for these analyses are diverse: Barnes' definition of exploratory talk is exclusively used in nine studies, Mercer's in two. Additionally, Mercer's coding scheme for social modes of thinking (CDE, cumulative - disputational - exploratory) is used in nine studies. Mercer emphasises that this coding scheme is, in the first place, a 'useful heuristic device [which helps] an analyst perceive the extent to which participants in a joint activity are at any stage (a) behaving cooperatively or competitevely and (b) engaging in the critical reflection or in the mutual acceptance of ideas' (Mercer, 2004, p. 146). In order to do so, the coding scheme zooms mostly in on key words in context, the length of utterances and other linguistic characteristics. Some characteristics of exploratory talk are analysed separately in 13 studies, e.g. explaining reasons, explaining disagreements and turn-taking (Kerawalla, 2015), turns, arguments, decision making strategies and group 'responsibility' (Polo et al., 2015), prosodical features (Skidmore & Murakami, 2010), etc. In 10 studies other frameworks related to exploratory talk are used. These are, for instance, the constant comparative method for student participation (Strauss & Corbin, 1990, cited in Maloch, 2002), Goffman's theory of footing (Skidmore and Murakami, 2010), four formats of classroom talk (Renshaw & Brown, 2007) and Fourlas' functional analysis system (Kumpulainen & Wray, 1999).

Quantitative studies

The two quantitative studies are statistical analyses on existing datasets, one of which focuses on aspects of exploratory talk and long term effects in a philosophy project (Topping & Trickey, 2014), while the other focuses on general learning outcomes in a whole-school approach where exploratory talk was imbedded in a learn-to-learn project (Mannion & Mercer, 2016).

Mixed methods studies

Of the 45 mixed methods studies Mercer's coding scheme for social modes of thinking (cumulative – disputational – exploratory) is used 23 times. Barnes' and Mercers' definitions of exploratory talk are used exclusively two times respectively one time. Specific characteristics of exploratory talk are analysed separately in 12 studies, e.g. shared participation, connection, reasoning and challenge (Cervetti, DiPardo & Staley, 2014), argument claims and levels of arguments (Nussbaum, 2005), descriptions, predictions, explanations, arguments (Webb & Treagust, 2006), etc. In 14 studies other frameworks related to exploratory talk are used. These are, for instance, models to analyse arguments (Beardsley, 1950; Toulmin, 1958, cited in Golanics & Nussbaum, 2008), to distinguish between transactive and non-transactive utterances (Kruger, 1993), to identify dimensions of classroom talk (Murcia & Sheffield, 2010), to analyse elaboration (Herrlitz-Biro et al., 2013) and to define group ownership (Haglund & Jeppsson, 2012).

In short, Mercer's coding scheme for social modes of thinking and his definition for exploratory talk (cf. previous chapter) are used most. Though it must be noted that about a quarter of these studies were carried out by researchers close to Mercer's research circle and by Mercer himself, the fact remains that only one quarter of the 88 studies uses a different analytical framework or combines Mercer's with other frameworks, choices which can be explained by specific research foci.

Experiment participation and education level

The 88 experimental research articles describe experiments in various school contexts and with pupils/students and/or teachers of different levels of education.

Table 10 gives an overview of the number of articles reporting on experiments involving pupils/students, teachers or both pupils/students and students. Distinction is also made for education level (kindergarten, primary, secondary, higher education). It must be said that in ten cases we found insufficient information to complete the table.

Table 10
Number of experimental studies based on participants' profile and educational level

	Kindergarten	**Primary education**	**Secondary education**	**Higher education**	**Total**
Pupils/students	3	41	9	11	64
Teachers		3	2	0	5
Pupils/students + teachers	1	9	0	1	11
Total	4	53	11	12	80

As table 10 shows, pupils/students were the focal participants in 64 experiments. 5 experiments focused mainly on teacher behaviour and in 11 experiments both pupils/students and one or more teachers were actively involved.

In the pupils/students experiments the talk of small groups of children/adolescents is analysed, either before and after an intervention and/or within a certain experimental context, e.g. problem solving tasks. Teacher experiments focus on the language of the teacher in whole-class or during small group approaches, the intervention mostly being a training in the use of exploratory talk during whole-class activities. Pupil/student-teacher experiments analyse the talk of both teacher and pupils/students in order to find correlations between both (e.g. the teacher models exploratory talk after which the effect on his pupils/students' talk is analysed).

Table 10 also shows that the majority of the experiments (57) took place in kindergarten and primary education, and a minority in secondary (11) and higher education (12). Specific reasons for this are seldom provided. There may be practical reasons (e.g. the fact that time-consuming interventions are easier to organise in primary than in secondary or higher education), but there may also be a strong pedagogical reason, as earlier research has shown that the amount and quality of small children's conversations are good predictors of educational attainment in secondary school (Hart & Risley, 1995).

Summarising, most experiments on exploratory talk have so far taken place in primary education and focused on pupil-pupil talk in small groups.

Sample size

Numbers about sample size in the experiments are not easy to provide. One reason is that in many articles no actual numbers are given. Another is that sample size of a target group not always coincides with the actual number of participants whose talk was analysed. Among the qualitative studies we found 14 experiments with six or less than participants whose talk was analysed. In the mixed methods studies we found only two target groups with six or less participants.

Interventions

Of the 88 experiments, 50 were interventional. Examples of such interventions are: teachers model exploratory talk for a limited period of time during which the talk of pupils is recorded and analysed afterwards; students are given various kinds of tasks (e.g. divergent vs. convergent tasks, book discussions, problem solving tasks) after which the exploratory nature of their talk is analysed; pupil talk is analysed before and after they have been taught the ground rules for exploratory talk, etc. The latter is by far the most common intervention.

Instruments

In most of the studies a variety of instruments is used to obtain data: collecting background information on school(s), teacher(s), pupil(s)/student(s); making field/ethnographic notes on educational activities, mostly during interventions; audio and/or videorecording of classroom talk; taking interviews from teacher(s) and/or pupil(s)/student(s); collecting written assignments, etc. Relevant conversations are transcribed for analysis, for which various coding schemes, conceptual frameworks and reflection scales are used, though mostly used is Mercer's CDE-scheme (cf. supra). Quantitative conversation analysis is done by means of concordancer software. When a problem solving test is included, almost exclusively Raven's Standard and/or Coloured Progressive Matrices[8] is used. Quantitative results are processed statistically, using appropriate (though seldom made explicit) software. In some cases, it is unclear which instruments were used.

---

[8] This test measures cognitive abilities of 6- to 60-year-olds without using language. It is a figural multiple-choice test, the taking of which takes approximately 40 minutes. The number of alternative answers is always the same. The test is free of educational and cultural bias (Raven, 1998).

3.2.3 Sociocultural discourse analysis

Within the variety of methods used in the 88 studies, one method stands out, i.e. sociocultural discourse analysis. In this paragraph we will discuss this method in detail. We will also zoom into the specific aspects of talk which can be measured in order to evaluate it as exploratory.

First, let us have another look at Mercer's definition of exploratory talk:

It is the kind of talk in which 'partners engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counterchallenged, but challenges are justified and alternative hypotheses are offered. […] Knowledge is made publicly accountable and reasoning is more visible in the talk. Progress then emerges from the eventual joint agreement reached' (Mercer, 1995, p. 104; 1996, p. 369; Wegerif & Mercer, 1997a, p. 53).

The difficulty with this definition (as with Barnes' and any other description) is how to make it operational. Through detailed conversation analysis – the bottom-up approach – a number of indicators for exploratory talk and also for disputational and cumulative talk may be obtained, but as said before a purely qualitative method has its limitations. If we want to generalise outcomes a top down, quantitative approach, using a number of pre-defined indicators, is required. Wegerif and Mercer (1997a) suggest four levels of analysis to capture the nature of types of talk in classroom conversations:

- Level 1 is about the fundamental way in which pupils orientate themselves towards each other when they start constructing knowledge. Which social mode of thinking do they employ: disputational, cumulative or exploratory?
- Level 2 focuses on the ground rules that govern the production of appropriate utterances, e.g. speech acts which are allowed and those which are not.
- Level 3 deals with specific speech acts or utterances classified according to their apparent function in the immediate context.
- Level 4 considers the actual, particular words recorded and transcribed.

Based on our literature study we want to slightly alter this structure and distinguish between level 4A (particular words recorded and transcribed) and an additional level 4B containing: turn-taking (Atwood et al., 2010; Kerawalla et al., 2013; Sutherland, 2013), long utterances (Dourneen, 2013; Mercer et al., 1999; Wegerif, 1996a) and quantity/quality of arguments (Nussbaum, 2005; Polo et al., 2015; Rojas-Drummond & Zapata, 2004; Webb, Nemer & Ing, 2006).

These levels represent levels of growing abstraction. Merely counting certain words (level 4), for instance, is more abstract than interpreting those words within a context (level 3). According to Wegerif and Mercer (1997a), in order to grasp all aspects of classroom talk explicit reference to these four levels is beneficial to explaining the nature and function of pupils' talk during collaborative activities. Also, this four-level analysis makes it possible to start off with a limited set of qualitative data which are analysed at a micro-level and then triangulated with a much larger set of quantitative data which are processed statistically and make generalisations possible (Wegerif, Mercer & Rojas-Drummond, 1999). Combining these types of data at different levels of analysis enables researchers to interpret results in a recursive and comprehensive way (Wegerif & Mercer, 1997a). This is what Wegerif and Mercer call the 'inverted dynamic pyramid' method. Since it was introduced in 1997, it has been used repeatedly by the same - and by other - researchers (cf. infra). As much research in this field focuses on (the effects of) exploratory talk, it is no surprise that blanks appear when browsing through literature in search of characteristic speech acts, words and utterances for the two other modes of thinking, cumulative and disputational talk.

Table 11 summarises the indicators of exploratory talk we found for all four levels, followed by further explanation. The sources of these indicators are listed in the last line of the table.

Table 11

Operationalised indicators for exploratory talk (4 levels)

Level 1		Level 2	Level 3	Level 4a	Level 4b
Social mode of thinking	**Main characteristics**	**Ground rules**	**Specific speech/communicative acts, exchanges**	**Language used: key words in context**	**Turn-taking, length of utterances, quantity and quality of arguments**
Exploratory	• active joint engagement with each others' ideas • alternative hypotheses offered • initiations challenged and counter-challenged • jusitifications given and developed • members try to collaborate and to understand each other's viewpoints • progress to joint acceptance of suggestions	1. all relevant information is shared 2. the group seeks to reach agreement 3. the group takes responsibility for decisions 4. reasons are expected 5. challenges are acceptable 6. alternatives are discussed before a	• assertion of knowledge • confirmation • critique • disagreement + explanation • explanation: procedural/conceptual • giving and asking for opinions • goal identification / clarification • hypothetical question • joint elaboration • justification of ideas • proposals + agreed action • proposals as an offer + explanation,	actually (to justify/clarify) agree also because (in reasoning) but (constructive challenging or clarification) could for example how idem I think/reckon/guess (introducing ideas) if (reason about problem)	• more long utterances • more symmetrical turn-taking • higher quantity and quality of arguments

		decision is taken 7. all in the group are encouraged to speak by other group members	clarification or elaboration • question seeking explanation, clarification or elaboration • reaching / stating a consensus • recounting another's idea • reformulating another's idea • selecting an option • self-regulation • utterances connect to the ones before	let's (cooperative suggestion) maybe (idem) might no (with justification or reasoning) what (idem) which/where (idem) why (task-related question) would you (in a question)	
Disputational	• disagreement without productive outcome • individualised decision-making • initiations lead to direct rejections or counter-propositions • members try to impose their own viewpoints		• disagreements • proposals + individual action • proposals as an imperative • short exchanges consisting of assertions and counter-assertions	no, it's not! yes, it is!	• mostly short utterances, • asymmetrical turn-taking, • low quantity and quality of arguments

	• no or little constructive criticism • no resolutions or resolutions not supported by agreement				
Cumulative	• accumulation • agreement without discussion • common knowledge constructed • members try to be friendly and to avoid conflicts • positive but uncritical adding of ideas • some effort to build consensus • superficial amendments that do not develop ideas	1. all relevant information is shared 2. the group seeks to reach agreement 3. the group takes responsibility for decisions	• joint proposals as offer or question • agreements • comments • confirmations • counter proposals • elaborations • questions seeking assistance, agreement or confirmation • repetitions	okay yes, that's good	• mostly short utterances, • turn-taking may be asymmetrical, • low quantity and quality of arguments
Sources	Dawes et al., 1992; Wegerif & Mercer, 1997; Nikolaidu, 2012; Fernandez et al., 2001	Wegerif & Mercer, 1997	Rojas-Drummond et al., 2006; Nikolaidu, 2012; Hewitt, 2014; Cervetti et al., 2016	Schmitz & Winskel, 2008; Kucirkova et al., 2014; Knight & Mercer, 2015; Mercer & Littleton, 2007; Boyd & Kong, 2015	Atwood et al., 2010; Kerawalla et al., 2013; Sutherland, 2013; Dourneen, 2013; Mercer et al., 1999; Wegerif, 1996a; Nussbaum, 2005; Polo et al., 2015; Rojas-Drummond & Zapata, 2004; Webb et al., 2006

Social modes of thinking (level 1)

It is important to note that social modes of thinking, i.e. exploratory, disputational and cumulative talk, are, above all, a way to understand the dynamics of social thinking (Wegerif, Mercer & Dawes, 1999). Wegerif and Mercer (1997a) apply Habermas' (1985) *Theory of Communicative Action* to argue that these modes describe fundamental orientations of participants in dialogue. Exploratory talk can be regarded as what Habermas calls 'communicative reality', i.e. 'reality defined […] through orientations and social ground-rules supporting a free and open encounter between ideas' (Wegerif, Mercer & Dawes, 1998, p. 200).

Ground rules (level 2)

As table 11 shows, seven ground rules of exploratory talk have been established. Wegerif (1996b) explains that ground rules are the implicit norms and expectations one has to take account of in order to participate successfully in educational discourse (Mercer & Hodgkinson, 2008). He refers to Sheeran and Barnes (1991) stating that few teachers are aware of such rules and, hence, rarely discuss them with their pupils. As a result, children often assemble their own heterogeneous set of ground rules, even if they have to receive them implicitly from TV chat shows (Hewitt, 1989). And if teachers do insist on certain requirements (e.g. 'Talk about it in groups', 'Discuss this and write down your conclusions'), they often do so without proper justification or without making sure children understand how they must talk together and to what purpose (Mercer, 1995; Mercer & Sams, 2006). Setati et al. (2002) add to this that learning from talk is significantly limited if it is not supported or complemented by strategies for learning to talk. Summarising, putting pupils together in small groups does not mean they will automatically use exploratory talk or collaborate smoothly (Wegerif, 2005). They need a framework, a set of relevant rules to hold on to. In short, they need ground rules.

Wegerif, Mercer and Dawes (1999) explain that the first three rules (all relevant information is shared, the group seeks to reach agreement and the group takes responsibility for decisions) are to bind the group, share together and construct knowledge together through seeking agreement. Rules four and five (reasons are expected, challenges are acceptable) are about explicit reasoning and challenging, two characteristics which distinguish exploratory talk from both other types. In exploratory talk challenges stimulate joint reasoning whereas in disputational talk they are all about 'winning' rather than about understanding or solving a problem together. In cumulative talk challenges are considered disruptive. The sixth rule (alternatives are discussed before a decision is taken) goes back to Kruger (1993) who found that groups that do best are those which consider alternatives before deciding. The seventh rule (all in the group are encouraged to speak by other group members) has been developed by Wegerif and Mercer, who found that pupils need to be actively encouraged by their peers to speak and to put forward ideas (Wegerif, Mercer & Dawes, 1999). According to Polo et al. (2015) the ground rules are consistent with the four indicators of productive argumentation of Asterhan: '[…] general willingness to listen and critically examine all the different ideas, willingness to make concessions in response to persuasive arguments, an atmosphere that is characterised by collaboration and mutual respect, and the activity is perceived as a competition between ideas, not between individuals' (Asterhan, 2013, p. 254, in Polo et al., 2015).

Literature does not reveal much discussion let alone disagreement about these ground rules. Gallon and Williamson (1992) were among the first to suggest that pupils must learn how to collaborate if collaboration is to be successful. Especially, it must be clear to them which type of discourse works best (Bullen et al., 2002). When used in other interventional studies the ground rules are used *as they are*, though often they are formulated with different words and listed in a different order (Bullen et al., 2002; Fernandez et al., 2001; Reninger & Rehark, 2009; Schmitz & Winskel, 2008). We found no substantial theoretical motives for these variations. In some cases the ground rules are rephrased to be better understood by e.g. small children. In some interventional studies they are the result of joint classroom debate, where pupils are introduced to exploratory talk in an inductive way and stimulated, through guided discovery (Bruner, 1962), to find and formulate the rules themselves. This explains that the set of zeven ground rules is occasionally extended to eight or nine rules or reduced to five or six.

Much experimental research on the use of exploratory talk includes interventions during which pupils are made aware of or just taught these ground rules, after which effects are measured such as the use of exploratory talk, problem solving skills, processes of identity-shaping, etc (Kerawalla, 2015; Mannion & Mercer, 2016; Mercer & Sams, 2006; Mercer et al., 1999; Rojas-Drummond et al., 2006; Rojas-Drummond et al., 2003; Rojas-Drummond & Zapata, 2004; Webb et al., 2016; Wegerif, 1996b; Wegerif, 2005; Wegerif & Mercer, 1997a; Wegerif et al., 1998). Some researchers explicitly acknowledge the necessity of introducing these ground rules and its added value for collaborating (Bullen et al., 2002; Luby, 2014; Maloch, 2002). Despite a large consensus about the need for pupils to apply these ground rules caution is advised by Polo et al. (2015), according to who ground rules 'do not systematically ban cumulative and disputational talk from group discussions, in the cases where they serve specific functions of wider, globally exploratory sequences.' (Polo et al., 2015, p. 29). Another warning comes from Rajala et al. (2012) whose research produced more insight on how pupils/students can successfully prompt each other's participation. Prompting questions such as 'What do you think?' and 'Why do you think that?' did not always function as expected, especially during *asymmetrical interaction* (i.e. group talk in which not every member can or is allowed to participate). This means that the use of prompts may have little effect on the acquisition of exploratory talk if the lack of symmetry is not dealt with, e.g. by intensifying feedback and feed forward on the talking process.

Turning those ground rules into measurable characteristics leaves a complex variety of operational indicators. Ground rules are about phrases, questions, analysis or inference, speculative or hypothetical talk (Sutherland, 2006, 2013), summarising words, tentativeness features, extended utterances, proportion of talk 'on task' (Janssen et al., 2010), relative equality of participation or turn-taking (Polo et al., 2015; Sutherland, 2006), about the quantity and quality or arguments (Rojas-Drummond & Zapata, 2004) and of course about the use of key words in context (Wegerif & Mercer, 1997a). They are also about face preservation and self-identity (Polo et al., 2015), learning ownership (Haglund & Jeppsson, 2012), metacognition (Mannion & Mercer, 2016), etc. It is no surprise that none of the 115 articles we studied deals with all these aspects simultaneously.

Speech acts (level 3)

Speech acts are characteristically performed in the utterance of sounds or the making or marks and intended to have meaning (Searle, 1969). Many studies publish lists of the speech/communicative acts they focus on. As table 11 shows, speech acts are not the utterances themselves but describe them. For instance, introducing an idea is a speech act, while 'I think' is the utterance by which it is realised. Exploratory talk is characterised by the presence of certain speech acts, e.g. asking for opinions, giving opinion, expressing disagreement with an explanation, posing hypothetical questions, etc. Some research has been done to very specific speech acts in the context of exploratory talk, e.g. adding and reacting (Tin, 2003), question-answer pairs (Atwood et al., 2010), language of possibility, reasoning links and pressing for reason (Boyd & Kong, 2017), reflective talk (Chick, 2015), generalisations and communicative struggles (Brown, 2016).

Words, phrases, utterances (level 4A)

Table 11 shows that, after qualitative analysis, some 20 words, phrases and utterances have been defined as indicators for exploratory talk, i.e. words that make explicit reasoning and ground rules like 'consensus building' or 'encouraging the other to formulate his opinion'. Most of them are conjunctions, modals and adverbials. Wegerif and Mercer (1997a) call them 'key words' but, as suggested earlier, it would be methodologically unsound to just count these key words in order to make statements about the exploratory nature of talk. Therefore Wegerif and Mercer (1997a) speak of 'key words in context' or of the 'key usage' of these words. Some researchers prefer the term 'reasoning words' or 'prompts for reasoning' (e.g. (Boyd & Kong, 2017; Soter et al., 2008; Sutherland, 2013), but it all comes down to this: conjunctions, modals and adverbials only become key words for exploratory talk or reasoning words when used in the appropriate context. Qualitative analysis must determine whether such words are indeed key words in context (i.e. code them as such) before any quantitative analyses can be done.

Another issue is the actual selection of key words. Languages differ from one another and so do people who participle at experiments. As explained earlier, in most studies about exploratory talk target groups are primary school children, but we also found studies which took place in kindergarten, secondary school, high school, at university and in experiments with groups of teachers. We believe it is feasible and pragmatic to make a selection of key words based on input from communicative grammars (e.g. Thomson and Martinet, 1986), translating the level 3 speech acts into concrete words and phrases, but literature provides no exhaustive list of age neutral let alone age bound key words. Actually, we found only a handful of studies which publish lists of 'their' key words. Schmitz and Winskel (2008), for instance, used the list from Littleton et al. (2005, also referring to Dawes, 2004; Dawes & Sams, 2004; Mercer et al., 2004; Mercer et al., 1999), but expanded it to fit into their study with older children.

All this strengthens the choice for a mixed methods approach: key words may be defined for usage in a coding scheme (top down), but the experimental context must remain flexible and give researchers the possibility to expand or alter the coding scheme based on qualitative analysis of talk (bottom up).

Other characteristics (Level 4B)

Though many studies and especially quantitative analyses focus on the use of key words (Rojas-Drummond & Mercer, 2003; Schmitz & Winskel, 2008; Sutherland, 2006; Wegerif, 1996a, 1996b; Wegerif, 2005; Wegerif & Mercer, 1997a; Wegerif, Mercer & Dawes, 1999; Wegerif, Mercer & Rojas-Drummond, 1999), other indicators of exploratory talk have been analysed as well. These are: turn-taking, length of utterances, the quantity and quality of arguments and – close to level 1: social modes of thinking – self-identity. We will now look at three of those indicators more closely.

Turn-taking

On the whole few studies deal with turn-taking as indicator of exploratory talk. Those that do, have limited themselves to a quantitative approach (counting terms and looking for participant (a)symmetries). This way, Littleton et al. (2005) and Wegerif (2005) found that the amount of asymmetrical talk diminished in most groups after an intervention in which pupils had learnt exploratory talking skills. The exact role of turn-taking remains a difficult issue, however. In their qualitative analyses Haglund and Jeppsson (2012) found rapid turn-taking to be a characteristic of exploratory talk, whereas Mercer (2000) describes smooth turn-taking as a characteristic of cumulative talk. In a qualitative study of university students' talk, Atwood et al. (2010) observed 'how clarity and coherence, as preconditions for learning, are achieved through turn-taking in exploratory talk within cooperative relations' (Atwood et al., 2010, p. 396). Nikolaidou (2012), on the other hand, found no relationship between turn-taking and the quality of talk whatsoever. The evidence of her study in primary school 'suggests clear relationships between the collaborative characteristics of peers' talk and action and the balance of collaboration; unfortunately, these characteristics cannot be captured by the turn-taking approach' (Nikolaidou, 2012, p. 762).

More clarity on the issue was provided by Rajala et al. (2012). They did a study on the relationship between exploratory talk and so-called asymmetries of talk in small groups of primary school children. They developed a framework to analyse turn-taking in exploratory talk in a very detailed way. In each discussion of their study task *quantitative (a)symmetry* was measured, simply by counting each participant's turn at talk. Next, the *interactional (a)symmetry* was analysed, i.e. the interactional power of a turn. This power is determined by combining two sets of indicators. The first set of indicators is *responsive* (the speaker relates to something that was said before) vs. *initiative* (the speaker is oriented to the future and creates the context for the next speakers). The second set of indicators describes whether a turn is *designated to the speaker*, *self-initiated* or whether it is *a stolen turn*. All indicators were again further divided into subcategories, ultimately leaving a framework that generated a value between 1 and 6 on an ordinal scale (the *IR index,* where *I = initiative* and *R = response*). 1 means that the turn has weak interactional power, 6 means it has very strong interactional power. Conversations with lots of short, stolen turns (low IR Index) are indicative for e.g. disputational talk, whereas conversations with strong responsive turns and e.g. many questions for clarification or elaboration are indicative for exploratory talk. These fine-grained analyses urged Rajala et al. (2012) to introduce two subcategories of exploratory talk, i.e. *inclusive* and *exclusive*, depending on the fact that all respectively not all participants were able to contribute to exploratory discussions (cf. supra).

The following transcript is illustrative. Two girls are talking about a moral dilemma, as Rajala et al. explain. In the story, a girl, Daniela, has forgotten her bag at school. Her mother has threatened to punish her if this is the case. The group has to decide whether Julia, Daniela's sister, should lie to her mother that she does not know about the bag. If she does, her sister is likely not to be punished.

Jenna:	Okay, so it reads here that should Julia lie to her mother that she doesn't know anything about the bag. (2.0)
Jenna:	*(Looks at Saana)* Should she or shouldn't she? (2.8)
Saana:	Too ridiculous. (1.2)
Jenna:	*(smiling)* So, she shouldn't? (1.8)
Saana:	Um (.) She shouldn't. Write that down. Or write down that she shouldn't.
Jenna:	*(Looks at Matti)* What do you think? (3.9)
Saana:	*(Looks at Matti)* Well? (2.2)
Matti:	*(smiling)* I don't [know.
Saana:	[I don't know.
Saana:	Yeah. Yeah, she should tell. So she gives her a spanking. *(smiling)* Yeah. (1.2) *(Jenna laughs)*
Matti:	I don't know.
Saana:	Or well. [Put there-
Jenna:	[or (.) shouldn't she?
Saana:	She should [tell (.) tell tell tell
Matti:	[She shouldn't.
Matti:	She shouldn't.
Saana:	She should. [Should should.
Jenna:	[Shouldn't.
Saana:	If you're of that opinion, *(smiling)* it's fine with me.

Figure 8. Illustrative fragment of asymmetric turn-taking (Rajala et al., 2012, p. 63)

Rajala et al. (2012) connect rapid and asymmetric turn-taking in this fragment with disputational talk. It is clear that Saana is showing evidence of interactive dominance. She takes nine turns, while Jaana takes six and Matti only four. Besides that, turns are very short and arguments are hardly formulated. Of course, this does not mean that the whole conversation should remain so. Immediately after this sequence, Jenna tries to reach a conclusion and invites the others to join her effort. Arguments are formulated and challenged, until a conclusion is reached. The conversation has become exploratory and turns take longer:

Jenna:	Okay, she shouldn't *(Looks at Matti)* Or, (0.5) she should (1.5) or should she or shouldn't she? (0.9)
Matti:	Well. *(Matti shrugs)*
Jenna:	I mean if she tells her little sister will get a spanking but if she doesn't tell then she won't. (1.0)
Matti:	Then she would be lying. (0.7)
Saana:	So she shouldn't lie to the mother, put down that she should giv- um put down that she gives her um that she should tell.
Jenna:	*(Looks at Saana)* Yeah, do we put down that she should tell? (0.6)
Saana:	Yeah.
Jenna:	*(Looks at Saana)* Do I put down that she should tell?
Saana:	Yeah.
Jenna:	Okay.

Figure 9. Illustrative fragment of symmetric turn-taking (Rajala et al., 2012, p. 63)

Long utterances

Several researchers found that pupils' length of utterances - made up of chains of clauses in order to provide evidence and reason - increased in post-intervention talk, i.e. after having mastered exploratory talk (Wegerif et al., 2005; Wegerif, 2005; Mercer et al., 2004; Mercer et al., 1999; Wegerif, Mercer & Dawes, 1999; Wegerif, Mercer & Rojas-Drummond, 1999). They suggest such longer utterances may be an indicator of increasing exploratory talk. In replication studies in Mexico the length of utterances was also taken into account, but a different measure had to be set. In the UK, researchers chose an arbitrary cut-off point of 100 characters in the transcripts, above which an utterance was considered to be 'long'. In Mexico that point had to be reduced to 70 in order to obtain sufficient material to analyse. It is conceivable that the length of utterances depends on language characteristics (in one language one needs more respectively less words to express the same thing than in the other). Another reason may have to do with educational tradition. As Wegerif et al. (2005) explain: 'Historically [in Mexican schools] there has been little interest in either group work or oracy. This is reflected both in the curriculum and in the physical layout of classrooms which all have desks in straight rows facing the front' (Wegerif et al., 2005, p. 40).

The quantity and quality of arguments

Rojas-Drummond et al. (2003) analysed the quantity and quality of arguments used during group talk. Specifically, they wanted to know whether the number of arguments used by pupils would increase after an intervention and whether the quality, i.e. the clarity and precision of those arguments would ameliorate. They distinguish four categories of argumentation, each with its linguistic characteristics: rudimentary arguments (mostly expressed by signalling and deixis), implicit arguments (all referents are implicit), semi-explicit arguments (congruent but unfinished statements, there is at least one non-explicit element) and explicit arguments (congruent and finished statements, all elements are explicit). Rojas-Drummond et al. (2003) found that children who learned exploratory talk not only produced more explicit arguments, the quality of these arguments increased as well. Combined with in-context key word count this means that exploratory talk has a visible positive impact on reasoning skills, i.e. on speech act level, above word level.

3.2.4 Conclusion and implications for Study 2

We started our methodological analysis in order to answer the following questions: how is exploratory talk being measured? Subquestions were: What kind of studies have been undertaken? Which research methods have been used? And which characteristics and variables of exploratory talk have been focused upon?

After selection we did a close reading of 88 studies involving empirical research and comparative analyses, methodology and theoretical issues. Of the empirical studies 41 were qualitative, two were quantitative and 45 used a mixed methods approach. 32 studies used Mercer's sociocultural discourse analysis. 57 empirical studies took place in kindergarten and primary education and focused on pupil-pupil talk, though we also found 11 studies which took teacher-pupil talk as their research focus, i.e. the extent to which teachers scaffold their pupils' talk (scaffolding being one of the keystones in Vygotsky's learning theory). 11 studies took place in secondary education and 12 in higher education.

Interventions included the teaching of ground rules of exploratory talk as defined by Mercer (1996) with pre- and post-test to measure effects. Though exploratory talk is considered to have learning effects across the curriculum, it is apparent that most studies chose science and mathematics as their experimental domain.

Further, literature shows that so far, a variety of characteristics of exploratory talk has been focused upon. On the one hand, this seems to fortify the impression that even after more than 20 years of increasing research there is still no complete clarity about the characteristics of exploratory talk. On the other hand, there seems to be a reasonable agreement about characteristics that cannot be ignored in analysis. These are key word usage, turn-taking, long utterances (elaboration) and the quality and quantity of arguments. As we believe the combined analysis of all characteristics provides a reasonable degree of certainty about the exploratory nature of talk, it is our intention to use these indicators in this dissertation (cf. Study 2) as well.

Considering the variety of analytical methods used we believe there is no real methodological consensus as far as measuring exploratory talk is concerned. In general, more than any other approach discursive analyses make an in-depth interpretative analysis possible, but if exploratory talk is to have any educational value, as is claimed by a growing number of researchers, we need generalising conclusions based on quantitative research as well. This is why we have decided to analyse data both qualitatively and quantitatively.

3.2.5 Discussion

As at the end of the previous chapter, we want to address some issues for discussions, which will be elaborated upon in Chapter 6, as indicated below.

- As problem solving should be part of a cross-curricular didactich approach, a problem solving test should appeal to more than scientific or mathematical thinking (Chapter 6, About some methodological issues – Research instruments: Raven's Progressive Matrices).
- Most studies measure to what extent a conversation is exploratory but not at what point it is or is not exploratory (Chapter 6, About studying exploratory talk).

3.3 Effects of exploratory talk (RQ 1.3)

In this section we want to discuss measured effects of exploratory talk. As pupils learn how to use exploratory talk, primarily they improve language skills which are inherent to this type of talk (cf. the definitions of Barnes, 1976, and Mercer, 1995), but other effects have also been noted, tested and demonstrated. Doing so, most interventional studies of exploratory talk share the same protocol: pupils and/or teachers are instructed how to talk exploratively via the use of ground rules (intervention). The results of pre- and post-tests on the one hand and in-depth analyses of exploratory conversations are then used to formulate effects. Based on our literature study and at the risk of imposing too rigid a structure, we wish to consider four main effect categories: linguistic, psychological, cognitive and pedagogical. Each of these categories can be operationalised, as is shown in table 12.

Table 12
Main effects of exploratory talk

Domain	Effects	Sources, illustrative
linguistic	more use of key words in context; use of more / better arguments; longer utterances; increased use of open-ended questions; better expression of opinions	Atwood et al., 2010; Boyd & Kong, 2017; Brown, 2016; Knight & Mercer, 2015; Kucirkova, Messer, Sheehy & Panadero, 2014; Kumpulainen & Wray, 1999; Mercer et al., 1999; Rojas-Drummond et al., 2003; Rojas-Drummond & Zapata, 2004; Webb et al., 2016; Wegerif & Mercer, 1997a; Wegerif, Mercer & Dawes, 1999
psychological	more equal involvement of pupils	Rajala et al., 2012; Rojas-Drummond et al., 2006
	pupils take up more assertive roles	Mercer, 1996; Wheeldon, 2006
	self-centredness is more replaced by group ownership; talk is more consensus driven	Enghag et al., 2009; Haglund & Jeppsson, 2012; Polo et al., 2015; Webb, 2015
	cultural divides are crossed	Tin, 2003
	stronger openness to external ideas	Dawes et al., 2010
	higher self-efficacy	Topping & Trickey, 2014; Webb et al., 2016
	increased growth mindset	Webb et al., 2016
cognitive	higher problem solving abilities	Mercer et al., 1999; Rojas-Drummond & Zapata, 2004; Topping & Trickey, 2014; Wegerif, 1996b

	more critical thinking	Soter et al., 2008
	better scientific and creative reasoning	Wegerif, Mercer & Dawes, 1999
	better verbal and non-verbal reasoning[9]	Topping & Trickey, 2014
	general learning gains academic performance	Luby, 2014; Mercer et al., 1999; Rajala et al., 2012; Tin, 2003
	internalisation of reasoning strategies on individual level	Wegerif, Mercer & Dawes, 1999
	better development of meaning and organisation of thought	Brevig, 2006; Golanics & Nussbaum, 2008; Webb et al., 2016
	increased confidence in writing	Robins, 2011
	stronger retrieval practice better memory	Webb et al., 2016
pedagogical	transfer to other domains	Mannion & Mercer, 2016; Riley, 2006; Webb et al., 2016
	better self-monitoring of learning and metacognitive reflection	Brevig, 2006; Mannion & Mercer, 2016
	positive attitude to undertake learning in group situations	Dawes et al., 2010
	reduction in dependency of the teacher	Gillies, 2014; Webb, 2015
	motivation for learning and persistence	Enghag et al., 2007; Mercer, 1996; Sutherland, 2013
	expanded Zone of Proximal Development	Fernandez et al., 2001; Rojas-Drummond & Mercer, 2003

Linguistic effects

It seems meaningful to look at the mastery of exploratory talk itself. In most experimental interventions the successful learning of exploratory talk is a condition which has to be fulfilled in order to generate other effects. It is therefore important to make certain that exploratory talk can actually be taught and learned, and that the teaching of exploratory talk can successfully transfer between educational contexts. The latter means that if exploratory talk is taught and learned in, say, sciences, its added value would be that pupils also employ it during group work for maths, languages, religion/ethics, world orientation and - in the end - in life outside the classroom. Quoting Galton and Williamson (1992) Mercer (1996) acknowledges that successful group work depends on the extent to which pupils are taught to collaborate. Ground rules, as we know from the previous section, are important means to teach children exploratory talk for learning in groups. It is obvious that research of effects of exploratory talk would have long come to a standstill if pupils had been unable to master this type of talk or mode of thinking. According to Mercer (1996) 'most children over 9 or 10 years old may have

[9] Reasoning is considered to be verbal as well as cognitive. The study of Topping and Trickey (2014) describes an intervention which involved Scottish 10-year-olds (n = 180). The target group was taught exploratory talk in the context of the Philosophy for Children (P4C) program. Changes in intervention classes included: increased use of open-ended questions by the teacher, increased participation of pupils in classroom dialog, and improved pupil reasoning in justification of opinions. The study also found gains in cognition based on the results of a Cognitive Abilities Test (Topping and Trickey, 2014).

all the language strategies they need to engage in exploratory talk (and so in educated discourse) without being expressly taught them. They may well already use them to good effect on occasion' (Mercer, 1996, p. 374). The question is, however, to what extent they use language as a social mode of thinking. '[…] we also need to consider what is the function and content of such conversations. Justifying social or moral choices to friends, or even discussing the social norms of classroom life (Much & Schweder, 1978; Elbers, 1994) is not necessarily the same as using language as a social mode of thinking when making joint decisions in solving problems or choosing between alternative explanations for observed physical events' (Mercer, 1996, p. 374). This is what a number studies successfully did as soon as exploratory talk was defined: finding out whether pupils can learn how to use it for learning (Mercer, 1996; Mercer et al., 1999; Wegerif, Mercer & Dawes, 1999, etc). Accordingly, by comparing pupil talk before and after introducing teachers to exploratory talk and dialogic teaching, the 'scaffolding' role of the teacher has repeatedly been measured (Barnes, 1976; Fernandez et al., 2001; Hunter, 2008; Maloch, 2002; Mercer & Sams, 2006; Murcia & Sheffield, 2010; Rojas-Drummond, Mercer & Dabrowski, 2001; Rojas-Drummond et al., 2003), but this falls beyond the scope of this dissertation.

Psychological effects

From the above, it seems obvious that effects of exploratory talk will first be visible on the linguistic level (or category), but as we know that exploratory talk, like disputational or cumulative talk, is a social mode of thinking (Mercer, 1996), effects will also and inevitably be psychological. E.g. when pupils work in dyads, triads or in groups of four or five, they will use language differently than Barnes' (1976) presentational talk, their interactions have the potential to become more symmetrical (Mercer & Sams, 2006), they can more easily develop group ownership while doing work (Haglund & Jeppsson, 2012), pupils take up more assertive roles (Mercer, 1996; Wheeldon, 2006), talk is more consensus driven (Tin, 2003), pupils are more open to ideas developed by others or found in other sources (Dawes, 2010), their feeling of self-efficacy increases (Topping & Trickey, 2014; Webb et al., 2016), collaborative attitudes like equity, active listening and goal-mindedness increase (Coultas, 2012; Golanics & Nussbaum, 2008; Kerr, 1998) and even cultural divides are more easily crossed (Tin, 2003).

Cognitive effects

A number of researchers have also tested cognitive effects in relationship with the use of exploratory talk, e.g. during problem solving activities. Apparently, such activities give pupils more opportunities to develop research skills (Mercer & Sams, 2006). As a matter of fact, Vygotsky claimed that solving problems in groups can enhance individual problem solving skills. The basic idea is that social - or intermental - activity can promote individual - or intramental - development, a process which is mediated by language (Mercer & Sams, 2006). Quite possibly, 'by internalizing or appropriating the ground rules […] pupils have come able to carry on a kind of silent rational dialogue with themselves' which explains gains in strategic thinking about problem issues (Rojas-Drummond & Mercer, 2003). In several empirical studies (e.g. Fernandez et al., 2001; Mercer, 1995; Mercer et al., 1999; Rojas-Drummond et al., 2006; Rojas-Drummond & Mercer, 2003; Webb et al., 2016; Wegerif & Mercer, 1997a, 1997b) researchers have therefore measured effects in problem solving skills after interventions. In order to measure these skills they mostly use Raven's (Coloured) Progressive Matrices (Raven, 1998). Apart from improving their problem solving skills, pupils also seem to become better at critical thinking and reasoning in general (Wegerif et al., 1999; Soter et al., 2008; Topping & Trickey, 2014), at the development of thought and meaning (Brevig, 2006; Golanics & Nussbaum, 2008; Webb et al., 2016) as well as at academic performance (Mercer et al., 1999; Tin, 2003; Rajala et al., 2012; Luby, 2014). Finally, Mercer (1996), Wegerif et al. (1999) and Rojas-Drummond (2003) have provided ample evidence for Vygotsky's claim that social learning precedes and stimulates individual learning.

Pedagogical effects

On a pedagogical level, positive effects of collaborative talk have been found during interventions comprising various school subjects, such as science (Cervetti et al., 2014; Dawes et al., 2010; Enghag et al., 2007; Haglund & Jeppsson, 2012; Polo et al., 2015), mathematics (Kassoti & Kliapis, 2009; Mercer & Sams, 2006; Murphy, 2015; Rojas-Drummond et al., 2001; Schmitz & Winskel, 2008; Webb, 2015), geography (Bullen et al., 2002), languages, i.e. reading comprehension, writing skills (Boyd & Kong, 2017; Brevig, 2006; Maloch, 2002; Soter et al., 2008), music education (Nikolaidou, 2012); reasoning skills in specific contexts such as philosophy for children (Topping & Trickey, 2014) and computer-based learning, including the use of iPads and the interactive whiteboard (Bowskill, 2010; Figueiredo, Figueiredo & Rego, 2015; Kucirkova et al., 2014; Murcia & Sheffield, 2010; Nikolaidou, 2012; Wegerif et al., 2003; Wegerif & Mercer, 1997a, 1997b); and also for the development of metacognitive skills (Brevig, 2006; Mannion & Mercer, 2016) and transfer to other contexts or communicative situations (Mannion & Mercer, 2016; Webb et al., 2016).

Though the boundaries between these four categories may not always be clear, we conclude that studies on exploratory talk show positive effects in several domains.

Chapter 4: Study 2: Empirical study – a quasi-experiment

RQ 1 of this dissertation was answered via a narrative review (Study 1). The results of that study will be taken into account in Study 2 – a quasi-experiment, which was set up in order to answer RQ 2-5.

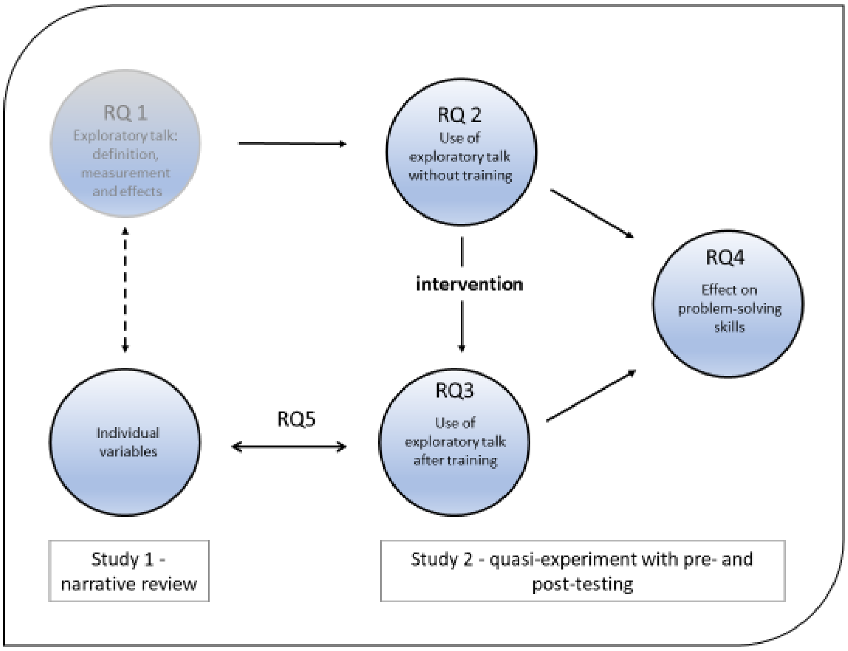


Figure 10. Global research design: focus on RQ 2-5

We will now first describe the research design (1), formulate hypotheses (2) and then discuss the methodology (3) and the results of the quasi-experiment (4), i.e. the pilot study (4.1) and the main study (4.2).

1. Research design

Exploratory talk and its educational potential have been the subject of international research and classroom practice, ever since Barnes launched the concept in 1976 and - a fortiori - since Mercer defined it based on empirical studies in 1995 and later. The last ten years research on the topic has deepened and widened, and become more international. As a tool for learning exploratory talk offers positive answers to teachers who question the successful implementation of collaborative activity (e.g. group work) and dialogic teaching and who doubt their importance for constructivist teaching. It is quite strange, then, that in Flemish education exploratory talk and its educational value seem virtually unknown, in classroom practice as well as in research. The main aim of this study is to change this and fill the gap.

Our literature study has shown that after more than 20 years of ever growing research and more critical approaches of the concept Mercer's (1995) definition of exploratory talk remains practically untouched. Based on our analysis of the measurement of exploratory talk we also believe that sociocultural discourse analysis, which combines qualitative and quantitative methods, will fit our purposes, adding the analysis of arguments as in Rojas-Drummond and Zapata (2004). As this study is the first of its kind in Flemish education, we believe it would be unwise to go ‘where no man has gone before’. On the contrary, it would be interesting to see if this study could validate the findings of Mercer (1996) and other researchers (e.g. Wegerif, 1999; Fernandez et al., 2001; Rojas-Drummond et al., 2004; Sutherland, 2006) in a new context, i.e. Flemish education, where we expect similar results. Therefore, we want to stay close to Mercer’s methodological approach.

To measure the learning effects of exploratory talks Mercer and his colleagues did various empirical experiments during which pupils first learned the basic rules of exploratory talks and were then given assignments involving exploratory talk (Mercer & Littleton, 2007). The experiments lasted several weeks. Through pre-and post-testing, the quantitative learning effects of exploratory talk were measured at group level and on an individual level. Qualitative analysis of interviews mapped the extent to which pupils mastered exploratory conversation as the experiment proceeded. As an example, we briefly describe 'Study 2: Effects on talk, reasoning and studying science and mathematics in Year 5’ (Mercer & Littleton, 2007). The research was conducted with pupils aged 9 and 10 years old. The focus was on talk quality during group work and group learning in the areas of learning science and mathematics.

406 children and 14 teachers from various primary schools took part in the research. The researchers worked with experimental groups (who went through the program) and control groups (who processed the same subject matter through group work but without receiving training in exploratory talk). The effect of learning on talking, reasoning and learning of the children was studied through observation and formal evaluation (using e.g. Raven’s Progressive Matrices) in experimental groups. The pupils from both the experimental and control group were submitted to a pre- and post-test. There were no specific criteria used to select the teachers for this experiment, except that they would be willing and eager to participate. Seven teachers and their classes participated in the experimental phase.

Data were collected and processed as follows:

- video recordings of the pupils during the implementation of the program;
- pre-and post-intervention filming of predefined activities of a group of pupils in each experimental group and control group;
- video recordings of teacher-led classroom moments in the experimental group as they trained their pupils in the use of exploratory talk;

- submission of both groups to the Raven's Progressive Matrices[10] test as pre-and post-test;
- submission to a knowledge and understanding test of math and science, both before and after the experiment.

A qualitative analysis of recorded talk (which was transcribed afterwards) was to show that children in the experimental group would talk gradually more in an exploratory manner with each other and use exploratory talk in a growing productive way during math and science activities. The analysis was also checked in a 'double blind' way[11] by two researchers who were not involved in the study but were familiar with the methodical analysis of classroom interaction. They were asked to: a) identify characteristics of exploratory talk in recorded conversations, b) note the extent to which such characteristics occurred, and c) determine whether these characteristics appeared before or after the 'training in exploratory talk' had taken place. The double-blind study showed that the research methodology was valid and reliable.

Also, a quantitative analysis of the conversations was made, i.e. key word count (words that show explorative features like 'because',' I think','I agree', etc). The count was made for group conversations before and after the program in both the experimental and control groups. The key words were statistically processed in order to reveal effects like the increase in the number of indicator words in the experimental group after the tutorial. Both the mathematics and science tests (before and after) and the scores on the Raven's test (before and after) were processed as well.

Replication of Mercer's study involved:

- organising a quasi-experiment in primary education;
- working with control and target classes in authentic school contexts;
- selecting triads in both control and target classes for close observation;
- teacher-led training of the target group classes in the use of exploratory talk through the introduction of ground rules and via basic lessons, followed by further group work training on a weekly basis;
- pre- and post-testing of pupils concerning a) the use of exploratory talk and b) problem solving skills (Raven's Progressive Matrices);
- data collection via pre- and post-test, and by video and sound recording of group work;

[10] This test measures cognitive abilities of 6- to 60-year-olds without using language. It is a figural multiple-choice test, the taking of which takes approximately 40 minutes. The number of alternative answers is always the same. The test is free of educational and cultural bias (Raven, 1998).
[11] Both researchers did not know whether conversations belonged to the control group or the target group.

- qualitative and quantitative analysis of pupils' conversations with a focus on the use of key words in context, as well as turn-taking and long utterances (Mercer et al., 1999).

Alterations involved:

- skipping maths and science tests as pre- and post-tests because in Flemish education, these subjects are not assessed in a centralised way[12];

Additions involved:

- analysis of the quantity and quality of arguments (Rojas-Drummond & Zapata, 2004) in order to obtain a more concrete view on the functional use of key words as this reflects the pupils' reasoning skills.

Apart from this, one extra research focus was added, based on Janssen et al.'s (2010) meta-analysis and their criticism that the interaction between pupils is often disregarded in studies like these. The narrative review we did revealed limited interest in the impact of individual variables on the use of exploratory talk. Holmes (1992) discussed gender differences (see also Wegerif, 2005; Nussbaum, 2005; Topping & Trickey, 2007a). Sutherland (2006, 2013) worked with low, middle and high achievers, while Robins (2011) focused on highly gifted children. Pupils' socioeconomic status was addressed in some studies (Wegerif et al., 1999; Webb et al., 2016). A number of studies specifically connected pre- and post-tests of problem solving skills with interventions during mathematics (Mercer, 1996; Mercer & Littleton, 2007; Mercer et al., 1999; Rojas-Drummond et al., 2006; Rojas-Drummond & Mercer, 2003; Schmitz & Winskel, 2008; Wegerif, Mercer & Dawes, 1999; Wheeldon, 2006), but to our knowledge, the match between individual maths skills and the use of exploratory talk has not been examined yet. Therefore, we will not only investigate the effects of interactive processes at group level but also look for intervening variables, such as mathematics skills, at the individual use of exploratory talk.

2. Hypotheses

Based on our research questions and on the findings of previous studies (e.g. Mercer, 1995, 2013; Mercer & Littleton, 2007; Mercer & Sams, 2006; Mercer et al., 1999; Rojas-Drummond & Mercer, 2003) we hypothesise the following:

---

[12] Contrary to many OECD countries Flemish education has no centralised assessment system for separate school subjects. Evaluation is the school's, which often means the teachers' responsibility. Standardised – though not centralised - assessments do take place on a regular basis in primary education (LVS Spelling and LVS Mathematics) and in both primary and secondary education (realisation of attainment targets), but results are used for diagnostic purposes respectively as indicators of the general quality of education.

For RQ 2: When pupils have not been trained in the use of exploratory talk, pupil-pupil talk will generally be cumulative and/or disputational. Characteristics of exploratory talk may occur but their use will be sporadical and unsystematic. After 12 weeks of group work without training no significant change is expected.

For RQ 3: After 12 weeks of training in the use of exploratory talk, including an extended period to practice it in group work, pupil-pupil talk will be more exploratory. Characteristics of exploratory talk will be applied regularly and systematically. We expect significant differences with those groups of pupils that have been trained. The target group will use more key words in context, turn-taking will be more democratic, more long utterances will appear and pupils will use more arguments during talk; also the quality of their arguments will have increased.

For RQ 4: After being trained in and having practised the use of exploratory talk pupils will have increased their problem solving skills, both at group level and individually.

For RQ 5: Evolutions in the use of exploratory talk at group level will also be visible at individual level, i.e. all group members will increase their use of key words for exploratory talk more or less equally. However, certain individual characteristics stimulate or inhibit the individual use of exploratory talk to a certain extent.

In the next section we will describe the methods used to answer the research questions and check our hypotheses.

3. Methodology

3.1 General outline

In order to answer our research questions, we chose a quasi-experimental study with pre- and post-test (Shadish, Cook & Campbell, 2002) with pupils of primary education in experimental and control groups, in five different schools in Antwerp. Control and target groups performed group-oriented activities for 12 weeks, at the rate of two activities per week. Apart from this, two weeks were scheduled for pre- and post-tests, specific observation and interviews.

We had several reasons to start this quasi-experiment with a pilot study in a small target group, and then proceed to the main study: we realised this experiment was rather complex, as it involved different teachers and schools, a demanding time scale, and a number of technical and organisational challenges (Dawes, 2004; Dawes & Sams, 2004). As such, a pilot study gave us the opportunity to 'try things out' on an organisational and technical level and to make any necessary alterations for the main study.

In order to avoid any kind of interference, we never worked with control groups and target groups in the same school simultaneously during the main study. In order to analyse the group discussions, research activities were recorded on video. Sound recordings were made as backups.

3.2 The quasi-experiment

It was our intention to organise a quasi-experiment in five different schools so as to obtain not only qualitative but also sufficient quantitative results. In order to compose our research population, we contacted 60 primary schools in the city and suburbs of Antwerp. These schools were selected in cooperation with Karel de Grote Hogeschool, which funded this experiment and provided a database of partner schools for internship. Karel de Grote Hogeschool gave us the credentials we needed to make this experiment possible. On the whole, 12 schools reacted positively. After an introductory briefing in these schools, six schools eventually decided to take part in the experiment. However, due to unforeseen staff changes, one school dropped out a few weeks before the start of the experiment, leaving us with five schools to work with: one school was located near the centre of Antwerp, the others were suburbian schools. We asked for control and target groups, i.e. third level classes (11- and 12-year-old children), to take part in the experiment for a period of two years. We expected the children to be old enough to collaborate effectively by the age of 11. It becomes more difficult when they are younger (Nikolaidou, 2012).

In school 1, the centre school, only one teacher took part. Her class was not a level 3 but a level 2 class (10-year-olds). Also, the profile of her pupils (many children with a lower socioeconomic background and poorer language skills) did not match that of the other pupils in the experiment. It would therefore be difficult to find a matching control group. All this made it relatively easy for us to decide to use this class for our pilot study. The main study, with five control and five target groups, would take place in schools 2-5. At the same time, in school 2, two teachers wanted to take part, giving us two classes instead of one. So, in total, we had 11 classes to work with (see table 13).

Table 13

Schools, classes, levels, school profile and number of pupils in each class

School	Class	Geographical context	Level/form	School size	Number of pupils	Experiment type	Experiment phase
1	1	centre	2/4	377	20	pilot study	1
2	2	suburb	3/5	236	17	main study	2
2	3	suburb	3/5	236	18	main study	2
2	4	suburb	3/5	236	17	main study	2
2	5	suburb	3/5	236	18	main study	2
3	6	suburb	3/6	233	24	main study	2
3	11	suburb	3/6	233	18	main study	2
4	7	suburb	3/5	308	19	main study	2
4	10	suburb	3/5	308	15	main study	2
5	8	suburb	3/6	380	17	main study	2
5	9	suburb	3/5	380	15	main study	2

From the teachers no specific foreknowledge was required, except for a professional openness to the topic and to the teaching methods used. At the start of each experimental period we asked the teachers to form groups of three pupils. The selection of these triads was based on the following characteristics: mixed gender; mixed level; positive assignment of pupils to a group, i.e. no coercion or negative choice; pupils have sufficient language skills; observed groups remain unchanged during the whole experiment (Dawes, 2004; Dawes & Sams, 2004). Empirical evidence for this approach comes from various studies: Heller and Hollabaugh (1992) found that groups of three pupils, who were chosen by the teacher so as to obtain groups with mixed abilities and mixed social structure functioned well (see also Enghag et al., 2007). Additional conclusions were drawn by Edwards and Westgate (1994) who found secure and stable groups to be a precondition for exploratory talk in order to learn more. Also, the fact that group members know one another well might contribute to this precondition (see also Enghag et al., 2009). Another argument to choose mixed level groups comes from a study involving undergraduate students by Golanics and Nussbaum (2008) who found that low-issue knowledge groups[13] engaged more in cumulative talk than mixed groups with high-issue knowledge groups.

We excluded from close observation: pupils with special learning needs or behavioural disorders and pupils who have insufficient (Dutch) language skills. As exploratory talk requires a (strong) critical attitude we also excluded friendship groups from observation (Mercer & Littleton, 2007). For the same reason, care was taken that the pupils: a) were not assigned any specific functions or responsibilities, as is the case in frequently used methods such as CLIM ('cooperative learning in multicultural groups', see Van den Branden & Van Gorp, 2000), and b) were not given any elaborative writing tasks for as long as the focus of the group work was on talk (Mercer, 2010b).

In each class three triads were selected to be video recorded. Both their pre- and post-test group work and group work they did during the intervention were recorded. The other triads took part in the same class activities, including the pre- and post-tests, but were not video recorded. Pre- and post-test data of all groups were included in the data analysis. The quasi experiment was planned as shown in table 14.

[13] In their study, Golanics and Nussbaum (2008) explored the improvement of argumentation in asynchronous, web-based discussions through goal instructions. These are statements at the end of a discussion prompt that indicate what students should achieve. They gave the students a survey in which they had to self-report how knowledgeable that they believed themselves to be about the discussion issue. Based on the outcomes of the survey low- and high-issue knowledge were used as independent variables.

Table 14

Planning of the quasi-experiment

Sep 2014 – Jan 2015	Feb-Jun 2015	Sep 2015-Feb 2016
Pilot group, school 1	target groups 2-3, school 2	
Control groups 1-2, school 2	control group 3, school 3	target group 4, school 3
	control group 4, school 4	target group 5, school 4
	control group 5, school 5	target group 6, school 5

Ideally, we would have worked with control groups first and target groups after. Limited funding, problematic time scales and demands made by the schools (which were only willing to allow control groups if there would be a target group as well) forced us to work in a more integrated way. This means that in some schools we worked with a control group while at the same time working with a target group in another school. We also decided to build in the possibility for a control group to switch to target group in a next experimental period ('switching replication', Shadish et al., 2002). This happened twice and simultaneously in school 2. The teachers of this school were only fully briefed about the pedagogy of exploratory talk after the control phase was finished. By including all schools in an experimental phase as well, even if it meant switching replication within the same year, we were sure to motivate them sufficiently to experience the benefits of the project and to keep them all aboard. Summarising, in total 163 unique pupils in 63 triads took part (see table 15).

Table 15

Control and target groups for each class and level, recorded resp. not recorded, number of pupils

School	Class	Target groups		Control groups		Form	Number of pupils
		Number of each recorded triad	Number of each non-recorded triad	Number of each recorded triad	Number of each non-record triad		
1	1	1*, 2*, 3*	4*, 5*, 6*	---	---	4	20
2	2			7, 8, 9	10, 11	5	17
2	3			12, 13,14	15, 16, 17	5	18
2	4	18, 19, 20	21, 22			5	17
2	5	23, 24, 25	26, 27, 28			5	18
3	6	33, 34, 35	29, 30, 31, 32			6	24
4	7			38, 40, 42	36, 37, 39, 41	5	19
5	8			43, 44, 45	46, 47	6	17
5	9	48, 49, 52	50, 51, 53			5	15
4	10	54, 55, 57	56, 58			5	15
3	11			59, 60, 64	61, 62, 63	6	18

**these groups were part of the pilot study*

3.3 Organisation of the quasi-experiment

Two weeks before each experiment we collected background information about the school, the teachers and their pupils. One week before and after each experiment we also organised pre- and post-tests and took interviews from the teachers. In between, for 12 weeks, the control and experimental triads did group work for at least ten minutes per session in the classroom. During the actual experiment we asked the teachers of the control groups to organise group work twice a week on fixed days, usually on a Tuesday and on a Thursday. For practical reasons, we recorded group work once a week in each class. In order to keep tabs on group work performed in our absence, we asked the teachers to keep a log in which they described when they did group work and what it was about.

As far as the content of the group work was concerned, we asked the teachers of the control group to give their pupils problem solving tasks of various kinds in various domains of the curriculum (languages, mathematics, world orientaton, religion). This means they did not receive lesson materials or subjects from the researcher. We did give them a predesigned checklist with directives (cf. appendix 4). With this checklist we wanted to make sure the pupils would have to work collaboratively and not just ‘work in groups’ (cf. Chapter 2, section 3).

The teachers of the target groups received the same information and directives. Also, they were briefed about exploratory talk and how to introduce this to their classes (cf. infra).

Research data were collected by taking a problem solving test and by organising a discussion excercise about a non-curricular topic as pre- and post-test which were all video recorded, and by collecting background information through semi-standard interviews from the teachers. Table 16 describes these research activities in detail.

Table 16

Data collection and planning

Research phase	Week	Research activity / intervention
1. Use of exploratory talk	1	Intake and briefing/training of the teachers
	2	Pre-tests, video recordings and interviews
2. Mastering exploratory talk	3-6	Gradual introduction of the pupils into exploratory talk (by the teacher), observations
	7-14	Observation and video recording of group work and of the conversations of the selected triads. The researcher attends all recorded moments but does not intervene, unless for technical reasons
3-4. Effects of mastering exploratory talk	15	Post-tests and video recordings
	16	Debriefing

We will now describe the research activities in more detail.

Mercer and Littleton (2007) describe basic lessons to teach exploratory talk (week 3-6), which we translated into Dutch (T'Sas, 2013), so they could be used in the pilot study as well. The classroom teacher gave these lessons, adapting some materials to bring them closer to the pupils' world of experience or foreknowledge, e.g. by showing pictures of Belgian instead of British soccer players or by having the children discuss the choice for a new school logo. As such learning goals remained unchanged, but pupils were more motivated to collaborate.

For all group work after the basic lessons (week 7-14) the teacher was free to choose the subject as well as the topic. We had asked the teacher to focus for ten minutes on pure talk, i.e. to keep writing tasks and manipulating lesson materials for later. This was necessary, as we wanted pupils to concentrate on (exploratory) talk and reasoning as much as possible. Puzzles, writing sheets, booklets, etc would only divert pupils from doing so.

The pupils were given talk cards or 'prompts' during the basic lessons, which made some of the ground rules more tangible and which helped the pupils to adhere to the rules. From week 4 on the prompts were replaced by a projection of or a poster with the ground rules. Each time group work started, the teacher would read these ground rules and ask the pupils which rules they wanted to focus on.

The teacher was asked to recapitulate subject matter from the basic lessons during other classroom activities the rest of the week, so as to deepen the pupils' understanding of them and to develop their exploratory talking skills. Concretely, after any group activity she would ask the pupils questions such as: 'Who asked why-questions?' or 'How did listening help you to do this task?' or 'What would you do next time to make sure every member of your group gives his/her opinion?'

Timing

Week 1: Intake and briefing

1. During an intake with the school director and/or an interview with the teacher involved, we collected information about the school (number of pupils, number of teachers, pedagogical project, global context), about the teacher (age, experience, didactic approaches, affinity with group work and with exploratory talk) and about the pupils (class size, age and gender, GOK-indicators, LVS-data for spelling and mathematics).

GOK indicators or 'indicatoren gelijke onderwijskansen' (equal opportunities for education) were introduced in 2001 on a federal level as indicators for extra individual support. Schools received additional financial means to organise education for pupils who meet criteria like low economic family status, limited language skills, etc. In the course of this study GOK was replaced by SES (socioeconomic status), sharing the same goals. In this study we will consequently refer to GOK. LVS or 'Leerlingvolgsysteem' comprises a number of tests, analytical instruments and pedagocial or didactic suggestions for technical reading skills, spelling and mathematics. They were developed for primary education between 1997 and 2006 (Dudal, 2000; 2001).

We also informed the school director and the teacher about the general purpose of the experiment, i.e. the observation of learning processes during group work. Detailed briefings were only given as soon as a target group was set up.

2. The teachers of each target group in each school were fully briefed about the experiment and their future role in it, which would include giving introductory lessons about exploratory talk (week 3-6), constructing problem solving tasks, coaching their pupils and giving them feed forward and feedback on their use of ground rules during additional group work (week 7-14). We also gave them a two hours workshop about exploratory talk and the relevant didactics to integrate it in their lessons. This workshop was organised one or two weeks before the start of each experiment.

3. The teachers of the control group in each school were briefed about the group work project, but no details were given about (the didactics of) exploratory talk. They were asked to organise group work as described above, from week 3 to 14.

4. Practical arrangements were made about the day and time of recorded group work (once a week), classroom arrangement, sharing of lesson plans with the researcher in order to receive proper feedback, log keeping, etc.

Week 2: Baseline assessment of pupils through pre-tests, video recordings and interviews

1. The pupils did a non-verbal non-cultural problem solving test called the Raven's Progressive Matrices test (Raven, 1998), in group and individually (cf. infra), with a three days interval. Group talk while doing the tests was video recorded.

2. Observations: recordings were made of three triads for each target and control group who discussed a non-curricular subject. Based on these recordings the extent to which pupils master the ground rules of exploratory talk was mapped (Mercer & Littleton, 2007). These ground rules are:

1. all relevant information is shared;
2. the group seeks to reach agreement;
3. the group takes responsibility for decisions;
4. reasons are expected;
5. challenges are acceptable;
6. alternatives are discussed before a decision is taken;
7. all in the group are encouraged to speak by other group members.

The ground rules were translated to Dutch (and called 'basisregels') and rephrased so that children would better understand them. As 'taking responsibility' (ground rule 3) proved difficult to make understandable for the pupils in the pilot study, ground rules 2 and 3 were joined and some examples of key words were added (table 17).

Table 17

Adapted version of the ground rules

Basisregels	English translation: ground rules
1. We moedigen elkaar aan om te spreken. ('Wat denk jij?' 'Vind jij ook dat…?')	1. We encourage one another to speak ('What do you think?', 'Do you agree with…?')
2. We hebben respect voor elk idee en elke mening.	2. We respect every idea or opinion.
3. We kijken en luisteren naar elkaar.	3. We watch and listen while somebody else is talking.
4. We geven uitleg en vragen argumenten. ('Waarom denk je dat?', Ik denk dat…')	4. We give explanations and ask for arguments for claims ('Why do you think that?', 'I think that…')
5. We leggen uit waarom we het eens of niet eens zijn met iemand ('Ik vind het goed, want …', 'Ik vind het niet goed, want …')	5. We challenge ideas using arguments ('I agree, because …'; 'I disagree, because …')
6. We werken samen aan een besluit dat iedereen goed vindt.	6. We want to reach a decision everybody agrees with.

The ground rules were printed on a large poster and attached to the front wall of the classroom (figure 11).

Figure 11. Scale model of the original ground rules poster in each target group classroom

Week 3-6: Introduction and observations

Target group: gradual introduction of techniques of exploratory talk by the classroom teacher

The teacher of the target group gave 5 introductory lessons on speaking, listening, asking questions, etc. The approach was inductive and experiential, and included observed group work as well as a gradual tilting from teacher sent to pupil driven activity (four phases of modelling instruction, cf. IJzendoorn, 1998). For these introductory lessons the basic material (http://thinkingtogether.educ.cam.ac.uk/resources) had already been translated and adapted (T'Sas, 2013).

The first introductory lesson is about the benefits of talking and listening, the second about how one expresses thoughts and formulates arguments, the third about the combination of these conversational skills in stories, transfering to conversation practice in the classroom (Dawes, 2008). In lesson 4 and 5 all this is deepened, with a focus on consensus building and devising the ground rules by the pupils themselves. During these lessons the pupils use so-called 'talk cards' or 'prompts' (cf. Appendix 1.3). These are explicitations of the ground rules ('Talk', 'Listen', 'What do you think?', 'Why …?', 'I don't agree, because'…) and can be used both as reminders and stimuli during group work, and as tools to practice the ground rules. Shortly after the introductory lessons, the prompts are discarded. After the last introductory lesson, a poster with the ground rules is hung up in the classroom (see figure 11), reminding the pupils of them and being useful to the teacher when giving feed foward or feedback. The teacher can, for instance, use the ground rules to ask questions such as: 'How could you see that X was really listening to you?', 'Why were you unable to reach a decision?' or 'How would you see to it that next time everyone gets a chance to give his opinion?'.

Each introductory lesson took about one lesson period of 40 minutes, but each aspect was reinforced the rest of the week and recapitulated during other activities. E.g. after a scientific activity pupils discussed how listening or asking questions had helped them to do the activity or solve the problem properly. From then onwards, the ground rules were used each week to stimulate and assess the use of exploratory talk in group work within several domains.

Control group: group activities without 'exploratory training'

Week 7-14: Observations and video recordings

Target group: group activitities with feed forward and feedback on the use of exploratory talk

From week 7 onwards pupils worked in groups for eight weeks at fixed times on subject matter of their curriculum, which had been adapted to include the use of exploratory talk. Language and environmental study were the most appropriate subjects, if only because these subjects include a lot of language use and activity compared to e.g. mathematics or science. It was also necessary to recapitulate certain aspects of exploratory talk during these lessons.

The teachers of the target groups were given the opportunity to use the checklist for problem solving activities and change their teaching materials as they saw fit, e.g. they could give their pupils certain assignments which demanded more group work and group thinking than is usually the case.

Control group: group activities

Evidently, the pupils of the control group were not briefed on the didactic approach of exploratory talk. They did similar discussion tasks in triads as the pupils of the target group, but without being trained in the use of exploratory talk. They were told that recordings were made in the context of scientific research about teaching.

Week 15: Impact assessment among pupils through post-tests

The pupils of both groups did the respective second half of Raven's (Coloured) Progressive Matrices (cf. supra, week 2). In both the target group and control group the same triads who discussed a non-curricular subject in week 1-2, did so now, but taking a different topic. These discussions werd video recorded.

Week 16: Debriefing

At the end of the experiment teachers were informed about the planning of the study, e.g. when to expect results and relevant information for their classroom practice.

3.4 Data collection and measurements

3.4.1 Data collection

The following data were gathered:

Background information about the school, teachers and pupils (independent variables) were collected via direct request and interviews:

- school: number of pupils, number of teachers, demographic profile, pedagogical project, number of special needs pupils;
- teacher: age, teaching experience, preferred didactic approach, affinity with collaborative learning and exploratory talk;
- pupils: level/form, age, gender, GOK indicators (special support), LVS score for spelling, LVS score for mathematics.

Pre- and post-intervention video recordings of predefined research activities of three triads were made in each target and control group. These activities were:

- all triads and individual pupils were submitted to the Raven's (Coloured) Progressive Matrices as pre- and post-test;
- all triads did a non-curricular discussion as pre- and post-test.

3.4.2 Instruments for data collection

Pre- and post-test: non-curricular discussion (ncd)

We video recorded a ten minutes conversation of the selected triads about a non-curricular subject. Recording started the moment the pupils got on task. During the pre- and post-test they discussed questions and propositions like ‘What would make our playground suitable for every pupil and which ideas deserve priority?’, ‘Rich people should give half of their money to poor people’ or ‘Shadows can have any colour’, ‘Facebook is not allowed to use my profile picture for publicity’ or ‘Would we still need money if we could just exchange possessions?’

Pre- and post-test: problem solving Raven’s Progressive Matrices (rpm)

The pupils’ problem solving skills were measured using Raven’s Standard/Coloured Progressive Matrices (Raven, 1998). The selected triads were video recorded for ten minutes while they were having a problem solving discussion based on this test.

Raven’s Standard Progressive Matrices comprises 5 groups of 12 items (A, B, C, D, E) of increasing difficulty (total: 60 items). The test is non-verbal and non-cultural. Each item is a sort of logic puzzle with one piece missing. Pupils have to complete the puzzle, choosing the right piece out of 6 or 8 possible pieces. Raven’s Coloured Progressive Matrices is the coloured version of groups A and B of Raven’s Standard Progressive Matrices to which an in-between group AB has been added (total: 18 items). An example of such an item is shown in Figure 12.

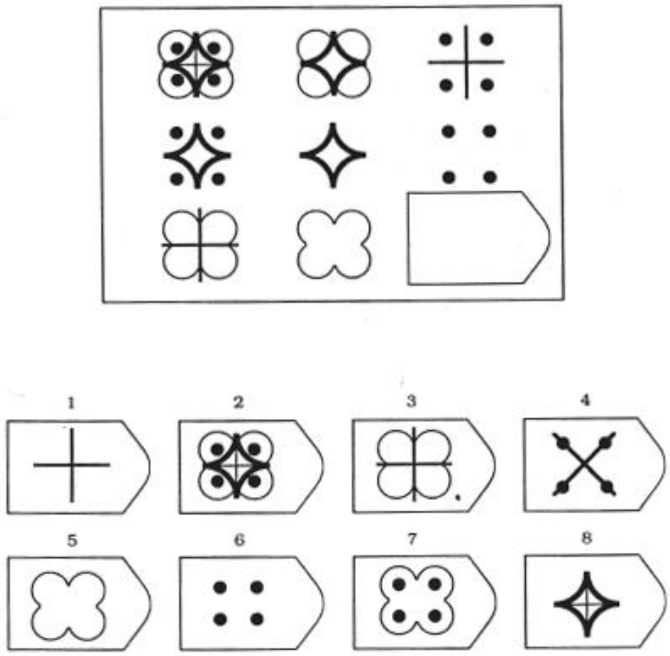


Figure 12. Illustrative item of Raven’s Progressive Matrices (Rojas-Drummond & Zapata, 2004)

In their research Mercer and Littleton (2007) account for the use of this test as follows:

‘Raven's Progressive Matrices [...] is commonly used as a general measure of non-verbal reasoning. It was designed to measure a person's ability to reason by analogy, independent their fiduciary language abilities and formal schooling. It is an excellent predictor of educational attainment and of occupational status.’ (Mercer & Littleton, 2007, p. 84)

In some of the earlier studies (e.g. Mercer et al., 1999; Rojas-Drummond & Zapata, 2004) for the group pre-test the pupils received the uneven test items, during the post-test groups received the even test. A few days later, for the individual pre-test the pupils received the coloured version of the even items, while receiving the uneven items during the post-test they received. In other studies (e.g. Rojas-Drummond et al., 2003) pupils were given the same version of the test twice, as pre-test and as post-test.

Since RQ 4 investigates effects at group level and individual level, the children were first submitted to this test in triads. Three to five days later, they did the individual test. For the pilot study the triad version of the test was Raven’s Standard Progressive Matrices: the pupils were given the even version of this test (18 items, A-, B- and C-level) as pre-test and the uneven version (idem) as post-test. The individual version was Raven’s Coloured Progressive Matrices (36 items, A-, AB- and C-level). This test was given to the pupils both as pre- and post-test.

Based ons our experiences in the pilot study, we changed the distribution pattern for the main study. This is discussed in detail in section 4.1.5.5. of this chapter.

Transcriptions

The recordings of target and control groups were transcribed by three students having at least a bachelor degree in languages. Transcriptions were kept simple: we asked for literal transcriptions on a lexical level (see Ochs, 1979). In order to enable automatic processing of the conversations with NVivo 10 dialect words or 'wrong' utterances were corrected to their standard equivalent, e.g. when a pupil said ['mɑ], meaning [ma:r], [ɣu:] meaning [ɣu:t], and also [ɑs] meaning [ɑls][14]. Thinking pauses and interruptions were indicated by specific symbols, like […] and [/]. If words or phrases were incomprehensible this was noted as [onverstaanbaar]. Non-relevant utterances such as repetitions, hm's or interjections were transcribed only if they had a communicative function (e.g. expressing surprise or agreement, indicating someone had not finished yet…). For practical reasons all this was done by the students but checked and corrected by the researcher afterwards.

3.5 Data processing

3.5.1 Sociocultural discourse analysis

For our analysis we applied Mercer's Inverted Dynamic Pyramid, i.e. a sociocultural discourse analysis (as explained in Chapter 3, section 3.2.3) which starts on the most abstract level and ends on the most concrete level and then returns to the abstract level, i.e. indicators for exploratory talk are selected by statistics and - when possible - statistics are drawn out from and informed by indicators for exploratory talk which were found via qualitative analysis. The method was first developed to compare the talk of groups of children in England and reported in Wegerif and Mercer (1997a). Essentially, this method combines qualitative and quantitave methods of research. Qualitatively, the way language is used in context is examined in detail by means of discourse analysis. Quantitatively, concordance software is used to analyse transcripts of larger amounts of language use. This way, specific examples of language use can be generalised without any loss of the context in which the language is used. In the method several levels of abstraction are integrated. In our research video recordings are the most concrete data. When pulling selected features out of these data, e.g. by making a transcription, the result is a first level of abstraction. The next level is pulling lists of key words in context out of the transcript. A subsequent level is doing a key word count using - in this case - NVivo 10 (Bazeley & Jackson, 2013), counting turn-taking and counting the length of utterances, and comparing these findings with similar data obtained from other transcripts. Whenever the key

[14] Both [ɑs] and [ɑls] are used meaning 'if', but taking both into account would complicate data processing, so we decided to use the most common form, i.e. [ɑls].

word count is finished, the researcher can go all the way back in order to qualitatively determine to what extent the key word context is actually contributing to exploratory talk and/or whether an argument is formulated. Iterative movements like these (from concrete to abstract and back to concrete, repeatedly if necessary) explain why the pyramid is called 'dynamic'.

As we have hypothesised, analysis of recorded talk (which was transcribed afterwards) would show that children in the target group talk gradually more in an exploratory manner with each other and use exploratory talk in a growing productive way during group work. Therefore, we analysed group talk of pre- and post-tests for all control and target groups.

Based on Study 1, the following elements were taken into account:

a. Key words in context

Following Dawes (2008), Mercer and Littleton (2007), Rojas-Drummond and Zapata (2004) and Wegerif and Mercer (1997a) we categorised key words in context and connected them to the ground rules used during the experiment (see table 18).

Table 18

Categories of key words in context

Category	Key words (Dutch)	Key words (English translation)	Ground rules
1	Words that indicate the organisation of ideas within each argument with conjunctions like 'omdat', 'want', 'en', 'daardoor', 'ook', 'als'… which are generally used to connect a claim or a challenge with one or more arguments and/or a conclusion.	Words that indicate the organisation of ideas within each argument with conjunctions like 'because (of)', 'and', 'also', 'if' … ' which are generally used to connect a claim or a challenge with one or more arguments and/or a conclusion.	2, 4
2	Affirmative and reinforcing utterances in connection with former claims, as 'Ik ga akkoord met…', 'Dat vind ik goed/oké', 'ja, dat is het'	Affirmative and reinforcing utterances in connection with former claims, as 'I agree with…', 'I like that', 'You are right', 'Yes, that's it' …	2, 3, 5
3	Questions pupils ask one another about the work they have to do, like 'Wat denk jij?', 'Wat vind jij (van)…', 'Denk je dat…', 'Geloof jij dat ook?'	Questions pupils ask one another about the work they have to do, like 'Wat do you think?', 'How about …?', 'Do you think that …?', 'Do you believe…?'	1

4	Why-questions combined with implicit and explicit reference to former claims, like ‘Waarom denk je dat (niet)?’	Why-questions combined with implicit and explicit reference to former claims, like ‘Why (do you say/think)…’	4
5	Concluding and summarising utterances like ‘Dus …’, ‘Onze conclusie is…’, ‘We beslissen dat …’, ‘We schrijven op dat …’, ‘Zijn we het dan eens over …?’, etc.	Concluding and summarising utterances like ‘So …’, ‘Our conclusion is…’, ‘We have decided that …’, ‘We write down …’, ‘So we all believe that…?’	6
6	Utterances that reflect opinion voicing and thinking processes, e.g. ‘Ik denk dat…’, ‘Ik geloof dat…’, ‘Volgens mij…’, ‘Ik vraag me af of …’, ‘Misschien…’, etc and modal auxiliaries, like ‘zou’, ‘kon’ …	Utterances that reflect opinion voicing and thinking processes, e.g. ‘I think/find that’, ‘wait…’, ‘I believe…’, ‘If you ask me …’, ‘I wonder if….’, ‘Perhaps …’, ‘Wouldn/Shouldn/Couldn’t we …’	4, 5

After the automated key word count via NVivo all discussions were thoroughly screened to rule out any of those key words of exploratory talk that were out of context (e.g. a pupil exclaiming ‘ja’ as a reaction to something said elsewhere in the classroom) or did not have a function in the exploratory conversation (e.g. a pupil shouting ‘I think it’s this’ as an expression of a choice but not as an indication of a thinking process). This analysis was checked by two researchers who were not involved in the study. In cases of doubt, mutual agreement was reached through discussion. For key word elimination Cohen’s kappa was 92, which denotes high degree of agreement[15].

b. Turn-taking

Transcripts of the problem solving (Raven’s Progressive Matrices) and non-curricular discussion were used to count the number of turns per pupil in each triad. The concept of quantitative (a)symmetry is determined by the presence or absence of more or less equal turn-taking during conversation, including the occurrence of interactive dominance/recession. When pupils talk disputationally in a group, they are more liable to interrupt one another, do little active listening and do not encourage each other, let alone more silent pupils, to voice their opinion (Mercer, 1995). We therefore expected to find more quantitative asymmetry than when they

[15] Cohen's kappa coefficient is a statistic which measures inter-rater agreement for qualitative (categorical) items. The kappa score can vary from 1 tot 0. Kappa assumes its theoretical maximum value of 1 only when both observers distribute codes the same. Very good agreement is indicated by a score between 1 and .80, good agreement by a score between .80 and .60, etc (Cohen, 1960).

talk exploratively. We also expected that some pupils would try to monopolise the conversation (interactive dominance), leaving little room to talk to more timid peers (interactive recession). The issue which we had to solve was where to draw the line between symmetrical and asymmetrical talk. Absolute symmetry would mean that every pupil in a triad takes one third of all turns; evidently, there would also be no interactive dominance or recession. Such symmetry is possible but not in real life. Furthermore, symmetry cannot only be quantitative but also qualitative (Rajala et al., 2012). As a qualitative study of (a)symmetry would enlarge the scope of this study beyond its limits, we decided to restrict it to a quantitative analysis. After discussing this with fellow researchers we decided a conversation would be quantitatively symmetrical when the difference between the number of turns of the most respectively least talking pupil is not higher than 12%. If higher than 12%, talk would be quantitatively asymmetrical. Further, we decided to speak of interactive dominance when one pupil of a triad took at least 45% of all turns and of interactive recession when less than 20%.

c. Length of utterances

Based on the transcripts of the problem solving and non-curricular discussions a count was made of all characters in utterances per pupil per triad in order to determine an evolution in length. As mentioned before 100 characters were chosen as cut-off point for long utterances (Mercer et al., 2004; Mercer et al., 1999; Wegerif et al., 2005; Wegerif, Mercer & Dawes, 1999; Wegerif, Mercer & Rojas-Drummond, 1999). This means that any utterance which consists of more than 100 characters is considered long. The longer utterances are, the higher the chance that we are dealing with elaborate reasoning, which is an indicator for exploratory talk. Rojas-Drummond and Zapata (2004), however, decided to reduce their cut-off point to 70 characters, because of the limited interest in either group work or oracy in Mexican education and, hence, the lack of a statistically usable amount of utterances longer than 100 characters.

It is difficult to determine what the cut-off point in this study should be. I-R-F studies in Flemish schools (Geudens et al., 1992; Van Gorp, 2010) suggest that interest in group work and oracy is rather limited as well, although attainment targets and curricula do promote these and so does teacher education. This might imply pupils would not be used to elaborating on their thoughts and hence produce mostly short utterances. Another issue is the difference in word length between languages. We could not find any study which compares word length in English and Dutch in discussions, but an English and Dutch translation of the same text from a third language, e.g. French, indeed shows a difference in the number of words or characters between both languages.[16] Finally, the nature of a text also plays a role in determining a cut-off

[16] Sources which revealed more about this issue were found in the translation and desktop publishing business world. Some translation agencies charge clients for each 'target' word (i.e. the number of words in the finished translation), others only charge for each 'source' word (i.e. the number of words in the original language). For clients who want to know how much a translation will cost them, this is valuable information. Likewise, desktop publishers want to know the length of translated text if they want to fit it into an already existing artwork file.

point for word length. Exchanging arguments in a discussion may well lead to longer utterances than doing so working on a puzzle, as this includes more non-verbal reasoning.

Since we did not find any conclusive evidence for expansion or contraction, we decided to make measurements for three different cut-off points in our study: 70, 100 and 115 characters, keeping in mind 100 characters as the main cut-off point to refer to.

d. Number and quality of arguments

Rojas-Drummond and Zapata (2004) support the claim made by Habermas (1985) that ‘learning processes by which we acquire knowledge of the world, by which we overcome our difficulties in comprehension and by which we renovate and extend our language are all supported in argumentation’ (Rojas-Drummond et al., 2004, p.543). They define argumentation as ‘the act of providings reasons to make admissible a certain position, opinion or conclusion, or to confront other’s positions, opinions or conclusions’ (Rojas-Drummond et al., 2004, p.540). An argument is then described as the combination of an assertion followed by one or more supporting utterances which can be several kinds of reasons. Both researchers worked out these concepts in a complete framework for argument building, which they used and which will also be used in this study as a tool to help analyse exploratory talk (table 19).

Lists of expansion and contraction published by international translation bureaus, like Kwintessential, Andiamo, Media Lingo and Linguasoft, suggest that an English text that is translated into German, which has many common features with Dutch, expands by 5 to 20% (Kwintessential) or even by 10 to 35% (Andiamo). Unfortunately, no comparison is made with Dutch. We did find figures about English vs. Dutch/Flemish text length on Media Lingo. Surprisingly, this company estimates an expansion of 10% for Flemish but a contraction of -5 tot -10% for Dutch. No explanation is given for this difference, but though there is a difference between the way the Flemish and the Dutch speak their common language (Standard Dutch), there is no such language as Standard Flemish. Linguasoft determines a -10% contraction when translating from English to Dutch, because of the used of compounds. But then again the use of compounds does not necessarily influence the number of characters in a substantial way.

Table 19

Categories and characteristics of arguments (Rojas-Drummond & Zapata, 2004)

Category	*Characteristic*	*Example*
Rudimentary arguments	Signalling and deixis	• Wait, yes, it's this one because look, it's like this. • No, because of this. • It would be this one, because look, it goes like this. • Ah, no, this one, yes, because look.
Implicit arguments	All referents are implicit	• It's this one, right? Yes, because look, it's in, it's in the middle and this one does not have it. • No, because look, it's separating and when it does, it goes together, yes? • Together, yes? • Yes, then it goes like this and it keeps moving until it touches, and it's on that side at the bottom, it's number 5
Semi-explicit arguments	Congruent but unfinished statements. There is at least one non-explicit element (underlined)	• No, the square / / the stripped square. There's one circle missing with <u>this one</u>. • It's the cross / / the second figure has the dots removed. Let's look at the sequence, <u>here</u> it has like this, they remove the X and the dots. It remains the cross. • We assume this figure has the dots removed and <u>here</u> it has <u>this</u> removed, then here an X would follow
Explicit arguments	Congruent and finished statements. All elements are explicit	• No, wait / / In the middle figure they remove only the little circles, it doesn't have dots, it would remain just the cross. • No, I say it's a white square, a square without the circle. Look, in the figure on the left they remove the circle, in the one at the bottom the diamond, and the square remains.

For reasons of interrater reliability, this analysis was checked by two researchers who were not involved in the study. In cases of doubt, mutual agreement was reached through discussion. For argument recognition Cohen's kappa was 72, which is fairly reliable. Table 20 illustrates how the analysis was done. We labelled the arguments with letters: A (rudimentary argument), B (implicit argument), C (semi-explicit argument) and D (explicit argument).

Table 20

Extract showing transcript and qualitative analysis for key words in context and (level of) arguments

Transcript problem solving, post-test, triad 3			Ground rule	Argument
Sara	D2.	D2.		
Peter	12.	12.		
Peter	Wie is er akkoord met dit plusje?	Who agrees with small plus?	1	
Sara	Alaya, ben jij akkoord met dit plusje?	Alaya, do you agree with this small plus?	1	
Alaya	*(wijst)* Dit, dit, dit.	*(points)* This, this, this.		
Sara	Ik ben akkoord met Peter en jij?	I agree with Peter and you?	6/1	
Alaya	Ja, ik ook. En waarom dat…	Yes, so do I. And why this …	6	
Peter	Waarom?	Why?	4	
Sara	Waarom denk je dat?	Why do you think it's that?	4	
Alaya	Omdat, hier… *(wijst)* hier er zo eentje, zo en een bolletje. Hier eh deze twee, maar niet deze dus die.	Because, here … *(points)* here's one, like this, and a small ball. Here eh these two, but not those.	5	B
Sara	Ik denk dat het kruisje hier moet, omdat ja… Dat staat hier *(wijst)* wel in, maar die veranderen heel de tijd van plaats.	I think there should be a cross here, because, yes… It does show right here *(points)*, but they change place all the time.	5	C
Peter	Ja, ik ben akkoord.	Yes, I agree	6	
Sara	Ben je akkoord? Waarom?	Do you agree? Why?	4	
Peter	*(wijst)* Omdat hier mist er eentje en en… Dat is… En overal, overal van dit is compleet en de bloemetjes zijn compleet, maar die plusjes zijn dus nog niet helemaal compleet. Dus kies ik voor plusje.	*(points)* Because there is one missing here and… It is… And everywhere, everywhere it is complete, the flowers are complete, but the pluses aren't. So I choose the small plus.	5	C

3.5.2 Data analysis: instruments

For the automated, quantitative analysis of transcriptions we needed specialised software, so-called concordancer software. As it is commonly used and supported in our university, we chose NVivo 10 (Bazeley & Jackson, 2013). This software marks predetermined linguistic elements in text, along with the partial context in which they occur. Examples of such linguistic elements or key words in context are: 'therefore', 'think', 'what?', 'why?', 'I think', etc (Dawes, 2008; Herrlitz-Biro et al., 2013; Mercer, 2010a, 2010b; Mercer & Littleton, 2007). A full list in Dutch can be found in Appendix 3. The count was made for the pre- and post-tests group conversations in both the experimental and control groups. Results were exported in Excel-format to be included in a dataset.

In order to prepare quantitative results for statistical processing standard spreadsheets were used. The following data were collected: quantitative (a)symmetry, the length of utterances, the number of key words and the results on the Raven's pre- and post-tests:

- Quantitative (a)symmetry has to do with the distribution of turns within a conversation. The more equal this distribution is, the more symmetrical the conversation becomes in terms of turn-taking. In order to determine quantitative (a)symmetry the number of turns each pupil of each triad takes, were counted.
- The length of utterances concerns all selected conversations, excluding text that is read aloud by pupil, e.g. the propositions used for the non-curricular discussion and also excluding off-talk utterances. As mentioned before (see Chapter 3, section 3.2.3), we chose an arbitrary cut-off point of 100 characters as the main referential cut-off point in the transcripts, above which utterances were considered to be 'long', but also considered 75 and 115 characters as cut-off points.
- The number of key words refers to all in-context key words used by each individual pupil in the selected triads. For this, a number of adjustments had to be made to the format of the transcriptions in Word, e.g. we had to replace pupils' names by unique numbers and replace non-relevant key words by nonsense words of the same length, lest they would be counted as well.
- The results on the Raven's pre- and post-tests and the data rendered by NVivo were processed both in order to construct a complete workable dataset for statistical processing.

Statistical processing (using SPSS 23) involved:

- frequency counts (number of markers for exploratory talk, number of arguments, group and individual scores for Raven's Progressive Matrices);
- significance analysis (differences between target and control groups concerning the use of markers, number of arguments, quantitative a/symmetry and interactive dominance/recession, length of utterances, group and individual scores for Raven's Progressive Matrices);
- correlation analysis of variables;
- mixed linear models analysis to determine the interaction of variables.

Due to insufficient information (LVS scores for spelling and mathematics were not available) statistical analyses of the influence of individual independent variables were not performed.

Summarising, table 21 gives an overview of the connection between the research questions, the collected data and the data analysis.

Table 21
Research, data and method of analysis

Research questions	Dataset	Method of analysis
1-2	background variables at start experiment	quantitative
	Raven's progressive matrices	quantitative
	video and sound recordings	quantitative + qualitative
3	video and sound recordings	quantitative + qualitative
4-5	Raven' s progressive matrices	quantitative
	transcripts	qualitative + quantitative

4. Results

Study 2 comprises two substudies: a pilot study and the main study. In this chapter we will first discuss the results of the pilot study (2.1). After that the results of the main study (2.2) will be presented.

4.1 Pilot study

We will first describe the background data obtained (4.1.1). Then we will discuss the results of the pre- and post-tests, i.e. characteristics of exploratory talk (4.1.2). After that the scores for Raven's (Coloured) Progressive Matrice at group and individual level (4.1.3) will be discussed. This means that, in our research paradigm, RQ 2, 3 and 4 are addressed:

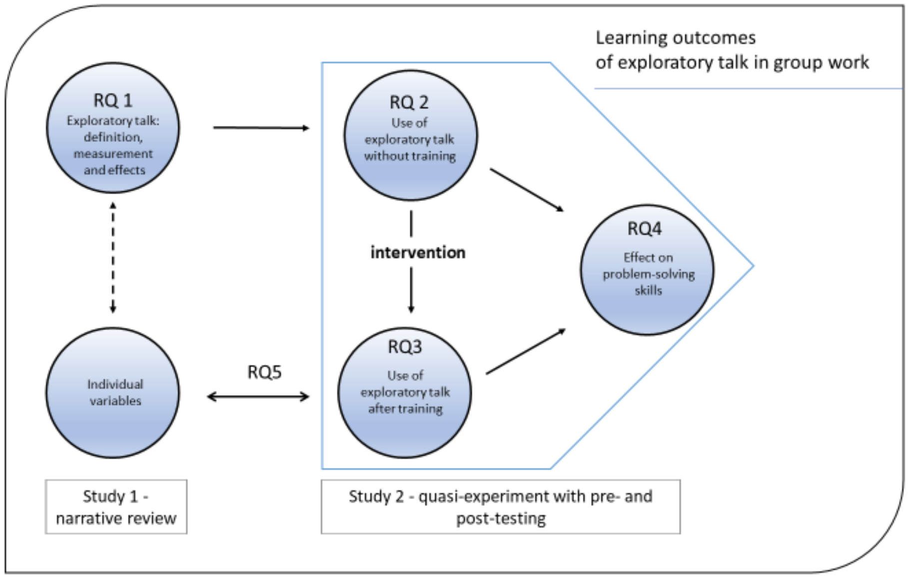


Figure 13. Research design of the pilot study

4.1.1 Background information

4.1.1.1 The school (school 1)

School 1 is an elementary school with kindergarten, which lies in a district close to the centre of Antwerp. The population in the school's neighbourhood is characterised by a social blend, including a considerable percentage of relatively young and socioeconomically weaker families. The school counts 377 pupils of 22 different nationalities. Of every six pupils, four have already repeated a year. This means that most pupils are older than is to be expected.

4.1.1.2 The teacher (teacher 1)

The class teacher (Teacher 1 in this study) is 33 years old and has 12 years of teaching experience. She received teacher education at Karel de Grote Hogeschool and had in-service training in various domains, e.g. coaching junior teachers. She finds it very important to give much structure to her pupils. Also, she combines a whole class approach with group work, individual work and other activating strategies. She regularly mixes groups of pupils and sometimes uses the CLIM method. At the start of the pilot study she had never heard of exploratory talk; the concept and underlying pedagogy were new to her.

4.1.1.3 The pupils (target group 1)

Table 22

Number of pupils in the control and target group, independent variables

Number	Group	Age	Gender	GOK
1	4	10	F	1
2	6	9	F	1
3	2	9	F	1
4	3	9	M	1
5	1	9	F	1
6	4	10	F	1
7	3	10	M	1
8	3	10	F	1
9	1	9	F	1
10	2	11	F	1
11	5	10	F	1
12	6	12	M	1
13	6	10	M	1
14	6	9	M	1
15	4	9	F	1
16	5	9	M	1
17	5	9	F	1
18	4	9	M	1
19	1	10	M	1
20	2	10	M	1

The pilot study class counts 20 pupils. They all live in the immediate neighbourhood of the school. Except for one they were all born in Belgium, but their origin is mostly foreign. The parents of 75% and the siblings of 50% of them do not speak Dutch as their mother tongue. Nine pupils are behind in their studies, i.e. they have had to repeat a year earlier in their school career or started later than usual. One pupil has special needs (GOK-coefficient 1,5), all are entitled to GOK support. Based on the testimony of the teacher, language skills are average to rather weak, e.g. the AVI-level (technical reading skills) of the pupils varies from 3 to 9 (9 being the highest level of the AVI-scale[17]).

[17] AVI is a system which allows primary schools to measure the progress of their pupils' technical reading skills. Pupils are AVI-tested once a year. Scores vary from 1 tot 9 (in the old system) and from 1 tot 12 in the new system which was implemented in 2013 and is gradually replacing the old system (Abimo, 2018; Boonen, 2006). Pupils are supposed to reach the maximum level by the age of 10.

4.1.2 Test results: exploratory talk

In this paragraph we will discuss the results of the pilot study, more specifically the results of the pre- and post-test regarding the use of key words (4.1.2.1), turn-taking (4.1.2.2), long utterances (4.1.2.3) and the use of arguments (4.1.2.4). But before doing that, we wish to make an illustrative analysis of two transcripts (translations from Dutch):

Transcript 1[18]

Pre-test: problem solving activity (Raven's Progressive Matrices)

Jamala:	Je zei het eerste.	Jamala:	You said the first one.
Zsusza:	Nee, ik zie B1. Het eerste… OK, B4.	Zsusza:	No, I said B1. The first one… OK, B4.
Pedro:	*(pakt het werkblad af van Zusza die het van tafel had genomen om het van dichter te kunnen bekijken)* Geef het aan mij. Geef het aan mij. Ik wil dit doen.	Pedro:	*(grabbing at the worksheet that Zusza has taken from the table in order to have a closer look)* Give it to me. Give it to me. I want to do this.
Jamala:	Nummer 2.	Jamala:	Number 2.
Zsusza:	Nee. Ja;	Zsusza:	No. Yes.
Pedro:	Ga terug.	Pedro:	Go back.
Zsusza:	Nee, het is nummer 2.	Zsusza:	No, it's number 2.
Pedro:	Ha nee, 2, 2. OK, sorry.	Pedro:	Ah no, 2, 2. OK, sorry.
Jamala:	Laat mij eens kijken.	Jamala:	Let me have a look.
Zsusza:	B6.	Zsusza:	B6.
Jamala:	Euh…	Jamala:	Eh…
Pedro:	O nee….	Pedro:	Oh no…
Jamala:	Nummer 3	Jamala:	Number 3.
Zsusza:	*(tegen Pedro, die van zijn stoel is gekomen en naast Zusza staat)* Nee, je kunt niet …	Zsusza:	*(to Pedro, who moves from his seat to stand next to Zsusza)* No, you can't...
Pedro:	Maar ik kan niets zien.	Pedro:	But I can't see a thing.
Jamala:	Je mag daar ook niet staan, nee.	Jamala:	You cannot stand here, no.
Pedro:	Mag niet, mag niet, ik kan niet blijven zitten.	Pedro:	Cannot, cannot, I cannot just sit.
Zsusza:	Nee, het mag niet.	Zsusza:	No, you can't.
Jamala:	Wie zegt dat? Wie zegt dat? Ha, nummer 5. Kom.	Jamala:	Says who? Says who? Ah, number 5. Come.
Zsusza:	Nee, het is nummer …	Zsusza:	No, it's number …
Jamala:	Maar het is nummer 5, niet?	Jamala:	But it is number 5, isn't it?
Pedro:	Nummmber 5.	Pedro:	Number 5.
Zsusza:	We zullen zien.	Zsusza:	We'll see about that.
Pedro:	Mij kan het niet schelen.	Pedro:	I don't care.

[18] All names are fictitious.

This extract very much illustrates how poorly the pupils collaborated during the pre-test: they interrupt each other, do little constructive reasoning, do not use many arguments (let alone elaborate on them) and at some point they just start quarreling. Further, turn-taking is chaotic and the atmosphere is competitive. In this fragment, for instance, Pedro first tries to take the work sheet out of Zsusza's hands and eventually even moves from his place to stand next to her, in order to get a clearer view of the worksheet. Also, key words of exploratory talk are rare or used in irrelevant contexts. There are a lot of 'no's' followed by unargumented claims, which indicates disputational talk. There is no building on each other's opinions, no knowledge whatsoever is shared and joint agreement is absent.

The following transcript illustrates a post-test conversation.

Transcript 2[19]
Post-test: problem solving (Raven's Progressive Matrices)

Imza:	Welke doen we, 6 of 3?	Imza:	Which one do we take, 6 or 3?
Eric:	3, ja.	Eric:	3, yes.
Berin:	Mag ik eens kijken?	Berin:	May I have a look?
Imza:	Ja, ik ben akkoord. Ik denk dat het drie is.	Imza:	Yes, I agree. I think it's 3.
Eric:	Mag ik de volgende doen?	Eric:	Can I do the next one?
Berin:	Ja, 3 is ook goed voor mij. Ja, jij…	Berin:	Yes, 3 for me, too. Yes you…
Eric:	Dankjewel, dat is heel vriendelijk.	Eric:	Thank you, that's very kind of you.
Berin:	Maar is het mijn beurt niet? Maar nee, geen probleem.	Berin:	Shouldn't I do this, oh no problem.
Eric:	Euh, welke… welke denk jij dat het is?	Eric:	Eh, which one… which one do you think it is?
Berin:	Euh…	Berin:	Eh…
Imza:	6. Ha nee.	Imza:	6. Ah no.
Eric:	*(wijst)* Ik denk dat het deze is.	Eric:	(*points)* I think it's this one.
Berin:	5, 5.	Berin:	5, 5.
Eric:	Nee, nee wacht.	Eric:	No, no wait…
Imza:	Of misschien 1.	Imza:	Or perhaps 1.
Eric:	Nee, nee, 4.	Eric:	No, no, 4.
Berin:	*(tegen Eric)* Toon het eens.	Berin:	*(to Eric)* Show us.
Eric:	4, weet je waarom? Omdat hier *(wijst)* een klein kruisje staat. Hier, zie je? Hier staat het. En dan zijn die er deze twee: vol, leeg, vol, leeg.	Eric:	4, do you know why? Because there's that little cross *(points)* here. Here, do you see it? Here it is. And then there are these two: full, empty, full, empty.
Imza:	Ja, je hebt gelijk, 4.	Imza:	Yes, you're right, 4.

[19] All names are fictitious.

Berin:	OK, 4. Iedereen akkoord?	Berin:	OK, 4. Does everyone agree?
Eric:	Ja.	Eric:	Yes.
Berin:	*(vult het gepaste vakje in op het werkblad)* OK.	Berin:	*(fills in the appropriate blank on the worksheet)* OK.

This dialogue is of a different kind. Pupils are much more collaborative. They listen to each other, interruptions are rare and turn-taking is democratic. In general, pupils build on propositions much more and before deciding what to fill in on the work sheet, they check if there is a consensus, asking if everyone agrees. They ask for an opinion as well as for a turn. Berin challenges Eric's claim that 'it is number 4' and when Eric explains his point, he explicitly starts his answer with a 'why-because' sentence. Finally, when Berin thinks it is her turn and starts to claim it, she stops her claim saying 'oh no problem'. By doing so she expresses priority of Polo et al.'s (2015) *group identity*, which can be indicative of exploratory talk (see also Wegerif et al., 2005). A number of key words for exploratory talk appear in the proper, constructive context: 'which one', 'why' / 'or perhaps' / '… agree?' / 'I think …'. Summarising, this extract is much more exploratory, as it contains visible examples of knowledge sharing, challenging, questioning and joint agreement.

Of course, extracts like these have little significance when it comes to formulating generalising results. They do, however, provide evidence for distinguishing types of talk (in these examples: disputational vs. exploratory) and for the quantitative analysis of indicators like in-context key word use. This is the subject of the next paragraph. By comparing these types of talk with the scores for the problem solving test, we were able to determine whether talk had a positive impact on the pupils' problem solving skills, a matter which will be discussed in 4.1.3.

4.1.2.1 Use of key words in context

In this paragraph we will first compare the use of key words for exploratory talk for three triads during the discussion of a non-curricular topic (ncd, 10'), before and after the intervention. Next, similar results will be presented as the triads were solving the Raven's Progressive Matrices (rpm), also during the pre- and post-test. Of each conversation ten minutes were analysed.

In table 23 we present an overview of the use of key words during the pre- and post-test non-curricular discussion. Each key word is allocated to one or more ground rules and to one of the six categories outlined in section 3.4.2.

Table 23

Key words use for during non-curricular discussion (ncd)

			Non-curricular discussion		
Category	**Ground rule**	**Key words**	**Pre-test**	**Post-test**	**Difference**
1	4,5	'als' [if]	28	37	9
1	4,5	'daarom' [therefore]	1	4	3
1	4,5	'en' [and]	104	143	39
1	4,5	'omdat' [because]	1	16	15
1	4,5	'ook' [also]	19	52	33
1	4,5	'want' [because]	10	53	43
2	2,4,6	'goe(d)' [good]	12	7	-5
2	2,4,6	'ja' [yes]	96	165	69
2	2,4,6	'oké' [okay]	24	22	-2
3	1	'vind jij' [do you find/think]	0	0	0
3	1	'wat denk jij' [what do you think]	0	2	2
3	1	'wat is' [what is]	5	33	28
4	4	'waarom' [why]	3	20	17
5	6	'akkoord' [agree(d)]	7	52	45
5	6	'besluit' [decide]	0	1	1
5	6	'dus' [so, thus]	11	47	36
6	4	'denk ik' [I think (*inversion*)]	1	1	0
6	4	'ik denk' [I think]	1	2	1
6	5	'maar dan' [but then (again)]	1	0	-1
6	5	'maar' [but]	15	95	80
6	4,5	'volgens mij' [as for me]	1	1	0
Total			340	753	413

Table 23 shows that pupils used a total of 340 key words doing the pre-test and 753 key words in the post-test. Based on a Wilcoxon Signed Rank test we can say the difference is significant (p = .001). The increase of key words is, however, not uniform. On the one hand 11 of the 22 key words are not used significantly more in the post-test than in the pre-test and some even slighlty less. On the other hand, key word increase is visible within each ground rule. Increases are substantial for 'ja', 'maar', 'want', 'akkoord', 'en', 'dus' ('yes', 'but', 'because', 'agree(d)', 'and', 'so/thus'). These key words are indicative for confirmation (GR 2-3), critical questions or challenges (GR 1-4-5), arguing (GR 1-5) and consensus building (GR 2-6).

In the next table (24), similar results as in table 23 are presented, but now we look at the pupils' conversations while they are doing the problem solving test before and after the intervention.

Table 24

Key word use during the problem solving test (Raven's Progressive Matrices - rpm)

			Raven's Progressive Matrices		
Category	**Ground rule**	**Key word**	**Pre-test**	**Post-test**	**Difference**
1	4,5	'als' [if]	3	12	9
1	4,5	'daarom' [therefore]	0	0	0
1	4,5	'en' [and]	41	56	15
1	4,5	'omdat' [because]	4	19	15
1	4,5	'ook' [also]	10	56	46
1	4,5	'want' [because]	25	40	15
2	2,4,6	'goe(d)' [good]	12	4	-8
2	2,4,6	'ja' [yes]	122	212	90
2	2,4,6	'oké' [okay]	27	41	14
3	1	'vind jij' [do you find/think]	0	0	0
3	1	'wat denk jij' [what do you think]	0	2	2
3	1	'wat is' [what is]	6	15	9
4	4	'waarom' [why]	3	30	27
5	6	'akkoord' [agree(d)]	0	131	131
5	6	'besluit' [decide]	0	0	0
5	6	'dus' [so, thus]	8	17	9
6	4	'denk ik' [I think (*inversion*)]	14	5	-9
6	4	'ik denk' [I think]	16	59	43
6	5	'maar dan' [but then (again)]	0	0	0
6	5	'maar' [but]	22	40	18
6	4	'volgens mij' [as for me]	0	0	0
Total			**313**	**739**	**426**

Table 24 shows that pupils used a total of 313 key words doing the pre-test and 739 key words in the post-test. The Wilcoxon Signed Rank Test demonstrated that the difference is significant (p = .001). Similar to the results for the non-curricular discussion 9 of the 22 key words are not used significantly more in the post-test than in the pre-test, but key word increase is visible within each ground rule. Increases are substantial for 'ja', 'ook', 'omdat', 'waarom', 'maar', 'ik denk', 'akkoord' ('yes', 'but', 'because', 'agree(d)', 'and', 'so/thus'). These key words are indicative for confirmation (GR 2-3), critical questions or challenges (GR 1-4-5), arguing (GR 1-5) and consensus building (GR 6).

Both tables, 23 and 24, indicate that key word use increases significantly, showing more organisation of thought, more critical questioning/challenging, arguing and consensus building.

4.1.2.2 Turn-taking

As pupils were mastering exploratory talk we expected them to interrupt one another less than in e.g. disputational talk and that they would exchange turns in a more democratic way (cf. ground rule 4: 'What do you think?'). We also expected pupils to be less dominant or less recessive during a conversation as group identity would become more important than self-identity (Polo et al., 2015). Therefore, we counted turns in each triad and calculated how these turns were distributed within the groups, before and after the intervention.

As explained in Chapter 4, section 3.5.1, we chose two arbitrary cut-off points: one to decide whether a conversation is a/symmetrical (less or more than 100 characters), another to determine interactive dominance/recession. Quantitative symmetry is established when the difference between the pupil taking the largest number of turns and the pupil taking the least turns is not more than 12% of all turns. Interactive dominance is determined when one pupil takes 45% or more of all turns, while interactive recession means a pupil takes less than 20% of all turns. Mathematically, this means that every conversation in which interaction is dominant or recessive, is asymmetrical by definition, but not all asymmetrical conversations need to include interactive dominance/recession.

Table 25 shows the percentages of turn-taking per pupil and per triad during problem solving conversations (Raven's Progressive Matrices) in the pre- and the post-test. It also shows the number of asymmetrical discussions (+/- 12%) and the number of instances of interactive dominance (+ 45%) or recession (- 20%).

Table 25
Turn-taking in problem solving conversations (rpm)

		Pre-test rpm			**Post-test rpm**		
Pupil	**Triad**	**% turn-taking**	**+/-12%**	**+45/-20%**	**% turn-taking**	**+/-12%**	**+45/-20%**
5	1	20.4			32.9		
9	1	40.8	1	0	33.3	0	0
13	1	38.8			33.8		
20	2	40.8			35.3		
3	2	28.6	1	0	34.5	0	0
10	2	30.5			30.2		
4	3	49.5			23.6		
8	3	3.9	1	1	41.4	1	0
7	3	46.6			34.9		
Total			**3**	**1**		**1**	**0**

Table 25 shows that during the pre-test, there is no quantitative symmetry in the discussions. The 12% margin for symmetry is exceeded by all triads. At the same time, interactive dominance and recession only occur in triad 3, in which one pupil hardly takes or receives a turn. This changes after the intervention. Interactive dominance and recession are no longer present in the post-test conversations, while in two of the three conversations turn-taking has become symmetrical.

Table 26 shows the percentages of turn-taking per pupil and per triad during non-curricular discussions for both pre- and post-tests.

Table 26
Turn-taking during the non-curricular discussion (ncd)

		Pre-test ncd			Post-test ncd		
Pupil	**Triad**	**% turn-taking**	**+/-12%**	**+45/-20%**	**% turn-taking**	**+/-12%**	**+45/-20%**
9	1	29.6			28.3		
13	1	32.4	0	0	34.3	0	0
5	1	38			37.4		
20	2	33.2			42.1		
3	2	31.7	0	0	36.5	0	0
10	2	35.1			21.3		
4	3	37.1			27.2		
8	3	31.3	0	0	35.4	0	0
7	3	31.6			37.4		
Total			**0**	**0**		**0**	**0**

The non-curricular discussions are characterised by quantitative symmetry in both pre- and post-test, as table 26 shows. Symmetry seems outspoken during the pre-test, as there are hardly any differences in turn-taking between the pupils. This is especially obvious in the discussion of triad 2. During the post-test, symmetry is less outspoken but turn-taking stays within symmetrical limits. Also, during both tests there is no interactive dominance or recession.

4.1.2.3 Length of utterances

Table 27 shows the number of long utterances (≥ 100 characters) and the length of the longest utterance for the three triads, with a distinction between the pre- and the post-test and for the non-curricular discussion (ncd) as well as for the problem solving test (rpm).

Table 27
Length of utterances in the non-curricular and problem solving discussions

Triad	**Pre-test**				**Post-test**			
	ncd		**rpm**		**ncd**		**rpm**	
	long utt.	**max**	**long utt.**	**max**	**long utt.**	**max**	**long utt.**	**max**
1	13	167	1	111	22	536	13	189
2	8	188	1	111	23	207	4	137
3	1	257	3	124	12	354	5	135

This table shows that the number of long utterances in the post-test increases in each triad compared to the pre-test. This is the most obvious in the non-curricular discussion (ncd). Based on a Wilcoxon Signed Rank test the difference is not significant, though (p = .109).

We wish to add that if we compare the longest utterance (max) in each conversation with one another similar increases are noted. Apparently, the intervention had a positive effect on the number of long utterances in all triads. Increase is outspoken in the non-curricular discussions.

4.1.2.4 Use of arguments

In order to analyse the quantity and quality of arguments used during the group discussions, we coded the arguments used by the pupils before and after the intervention. Coding was done separately by two researchers (Cohen's kappa =72, i.e. good reliability). The quality of these arguments was determined by mutual agreement (inter rater reliability), using the typology of Rojas-Drummond and Zapata (2004), see section 3.5.1, where A stands for a rudimentary argument, B for an implicit argument, C for a semi-explicit argument and D for an explicit argument.

Table 28 shows the number of each level of arguments and the total number of arguments during the non-curricular discussions (ncd) and the problem solving conversations with Raven's Progressive Matrices (rpm). The results of the three triads are mentioned, for the pre- as well as for the post-test.

Table 28
Number of arguments and levels of argumentation in the non-curricular and problem solving discussions

		Level of argumentation				
Activity	**Triad**	**LEV A**	**LEV B**	**LEV C**	**LEV D**	**Total**
Pre-test ncd	1	1	4	1		6
Post-test ncd	1		2	6	10	18
Pre-test rpm	1	4	6			11
Post-test rpm	1	5	11	1	1	18
Pre-test ncd	2	7	8			15
Post-test ncd	2	7	11	1		19
Pre-test rpm	2		7	4		11
Post-test rpm	2			3	13	16
Pre-test ncd	3	1	8			9
Post-test ncd	3	10	21	2		33
Pre-test rpm	3			4		4
Post-test rpm	3		4	12	6	22

In order to fully understand the evolution of the use of arguments, table 29 shows the average number of argumentation in all triads for the non-curriculuar discussion (ncd) and the problem solving discussion (rpm).

Table 29
Use of arguments – Average level of argumentation

		Average level of argumentation				
Discussion	**Measurement**	**LEV A**	**LEV B**	**LEV C**	**LEV D**	**Total**
Non-curricular	Pre-test	3	6.7	1	0	10
	Post-test	8.5	11.3	3	10	23.3
Problem solving	Pre-test	4	6.5	4	0	8.7
	Post-test	5	7.5	5.3	6.7	18.7

Both tables, 28 and 29, show that before the intervention, pupils do not use many arguments and if they do, the quality of these arguments is rather low (many A's and B's, i.e. many implicit and semi-implicit arguments). Higher level arguments (semi-explicit and explicit, level C and D) are rare.

After the intervention there is a significant increase of arguments, from an average of 10 in the pre-test to 23.3 in the post-test (Wilcoxon Signed Rank Test; p = .043) for the non-curricular discussion; and from an average of 8.7 in the pre-test to 18.7 in the post-test for the problem solving activity (Wilcoxon Signed Rank Test; p = .042). The pupils' quality of those arguments has also risen, as in the post-tests arguments of level C and D appear, which were almost absent in the pre-test.

When we compare the results of both conversations, it becomes apparent that the quality of arguments increases not as much during the problem solving test (rpm) than during the non-curricular discussion (ncd). This is not surprising, as during the discussion the pupils are not working with learning materials they can refer to. Hence, they have to be very explicit when they are expressing their thoughts or voicing their opinion. When solving Raven's Progressive Matrices, pupils are bound to talk more implicitly as they can 'replace' utterances by pointing at the puzzles and use the test sheets to demonstrate what they mean. On the whole, it is clear that the intervention has stimulated the pupils to use more arguments and to be more explicit when formulating them.

4.1.3 Test results: problem solving skills

In this paragraph we will discuss the problem solving skills of the pupils (scores for Raven's Standard/Coloured Progressive Matrices as pre- and post-test).

4.1.3.1 Scores at group level

In order to measure the problem solving skills of the pupils at group level Raven's Standard Progressive Matrices was used. The pupils did a carefully selected half[20] of the items of the test as pre-test and the other half as post-test. Each test comprised 18 items. The results of the groups' score for Raven's Standard Progressive Matrices (A, B and C level) on the pre-test (uneven) and the post-test (even) for the six triads are shown in table 30.

Table 30
Groups' score for Raven's Standard Progressive Matrices (rpm)

		Score pre-test	Score post-test
N	**Valid**	6	6
	Missing	0	0
Mean		12.50	16.33
Std. Deviation		2.345	.816
Range		6	2
Minimum		11	15
Maximum		17	17

As table 30 shows, the average score of the six groups that did the test is 12.50 on a total of 18 on the pre-test and 16.33 to 18 on the post-test, which is a significant increase by 30.6% (Wilcoxon Signed Rank test; $p = .045$). The target group seems to have strongly benefited from the intervention, as it has improved its problem solving skills in a substantial way.

4.1.3.2 Scores at individual level

In order to measure the problem solving skills of the pupils at individual level Raven's Coloured Progressive Matrices was used. The test comprised 36 items and was done twice, as pre-test and as post-test, as in Wegerif (1996b). The following table shows the results of the individual scores for Raven's Coloured Progressive Matrices on a total of 36 points, for the pre- as well as for the post-test.

[20] For the selection of the items two criteria were used: a) the progressive difficulty of the items had to be respected and b) the level of difficulty of both pre- and post-tests had to be equal. For a full description, see Chapter 3, paragraph 3.2.4.

Table 31
Individual scores for Raven's Coloured Progressive Matrices (rpm)

	Individual score pre-test	Individual score post-test
N Valid	19	18
Missing	1	2
Mean	28.11	31.06
Std. Deviation	5.517	3.506
Range	22	14
Minimum	13	22
Maximum	35	36

Individually, pupils score an average of 28.11 out of 36 for the pre-test and 31.06 out of 36 for the post-test. This is a significant average increase of 2.95 points or 10.5% (Wilcoxon Signed Rank Test; $p = .001$). Comparatively, in a similar study with the same method of individual re-testing, Mercer et al. (1999) saw target group pupils increase their scores by 10% after an eight week intervention.

4.1.4 Tentative conclusions

Based on the outcomes of the pilot study we can draw some tentative conclusions regarding the use of exploratory talk before and after the intervention, and about the effect on the pupils' problem solving skills.

4.1.4.1 Use of exploratory talk in group assignments

Pre-test scores, observations of group work, qualitative analysis of the transcripts and quantitative analyses of key word use, turn-taking, long utterances and use of arguments point out that the pupils of the pilot study do not use exploratory talk systematically before the start of the experiment. Their key word use in context is low and so are long utterances which would suggest elaboration. Why-questions and utterances such as 'ik denk', 'omdat/want' of 'akkoord' ('I think', 'because', 'agree(d)'), which indicate thought processes, argumentation and consensus building, are rare. Pupils interrupt one another regularly, divert from the task or topic, do not systematically distribute turns in a democratic way and show evidence of interactive dominance/recession. Further, they do not use many arguments and if they do, the arguments are mostly implicit and semi-implicit. We conclude that the pupils do not master the ground rules for exploratory talk.

Post-test analyses show that the language of pupils has been positively influenced by the intervention: key word use and the number of long utterances have increased. There is more quantitative symmetry in turn-taking, while interactive dominance or recession is completely absent. The number and quality of arguments have also increased significantly. It seems that in general, pupils' reasoning skills have improved. We conclude that pupils use exploratory talk in a consistent and systematic way.

These results are very interesting, as these are pupils with lower language skills. In their study with pupils with a similar profile, Herrlitz-Biro et al. (2013) found that pupils 'sometimes used key words but that this was not per se an indicator of high-quality reasoning' (Herrlitz-Biro et al., 2013, p. 1409). In their study they also found that pupils sometimes did not use key words because they did not speak any Dutch, the language in which they were taught. We did notice that some of the pupils in our pilot study used key words in the wrong way, e.g. saying 'akkoord' ('I agree') not as a consensus builder but as an expression of choice between possibilities in a task. However, we did not observe a lack of using key words nor a lack of high-order reasoning, cf. the increasing number and quality of arguments. But then again, in the Herrlitz-Biro et al. (2013) study some pupils did not seem to speak any Dutch at all, whereas the pupils in our study did, though on a lower level. Interpreting our results, we suggest that language-weak pupils start using some of the key words as new vocabulary which they employ to organise their thoughts. Apparently, they have learnt to use key words which they did not seem to use in the same context or even did not use at all before the intervention. As our observations show, especially during the problem solving discussions pupils used much non-verbal language and implicit arguments. By improving their reasoning skills, non-verbal language has become less necessary.

Another interesting issue is that pupils have improved symmetrical turn-taking during the problem solving conversation, whereas no significant evolution shows for the non-curricular discussion. We believe this can be explained by the fact that the Raven's problem solving test evokes more right-or-wrong discussions, taking the discussion closer to a disputational type of talk. Analysis of the non-curricular discussions also showed lower use of arguments and more cumulative talk during the pre-test. As cumulative talk is characterised by the need to preserve a safe group atmosphere (Polo et al., 2015), it is conceivable that, especially in the pre-test, pupils unconsciously did their best to avoid interactive dominance and quantitative asymmetry. As pupils learn to talk exploratively, the non-curricular discussions became more vivid, which may explain the observed tendency towards less symmetry.

4.1.4.2 Effects on pupils' problem solving skills at group and at individual level

The results of the pilot study are similar to the results of previously mentioned research (Fernandez et al., 2001; Mercer et al., 1999; Rojas-Drummond et al., 2006; Rojas-Drummond & Mercer, 2003; Rojas-Drummond & Zapata, 2004; Wegerif, 1996b; Wegerif, 2005; Wegerif & Mercer, 1997a): problem solving skills of pupils increase significantly after the intervention, both at group and individual level. This goes hand in hand with the – also significant – increase of the use of key words for exploratory talk, more democratic turn-taking, more long utterances and the increased use of arguments and their quality. This means that the type of talk pupils use has had an effect on their problem solving skills. These results reflect Vygotsky's theory that learning first takes place in a social context (intermental) and is then taken up by the individual (intramental). In other words, by learning exploratory talk and by using it for collaborative work the children have developed reasoning strategies in group. All in line with Vygotsky they have then integrated this new knowledge and those skills at an individual level.

4.1.4.3 Individual variables

As mentioned before, we did not have sufficient data to answer RQ 5, but we would like to compare our preliminary findings with Sutherland (2013). After having 12- and 13-year-old pupils taught the ground rules of exploratory talk and reflexive skills, Sutherland (2006) noted 'a significant reversal in attitude for some students, especially boys in classes of lower socioeconomic status, moving from scepticism or dislike of group talk to valuing it for learning' (Sutherland, 2013, p. 20). Though our observations and interviews did not reveal any dislike for group work, the recorded images seemed to suggest a similar positive attitudinal shift: before the intervention most group work was rather chaotic and competitive. It resembled Polo et al.'s (2015) 'competitive footing', which is marked by a focus on self-identity and face-preserving talk. We believe it evolved to more 'constructively-critical footing', which is marked by a focus on group-identity. Pupils raised their voices less than before, they remained more on task, listened to one another and if their was a problem to be solved, they took this as a joint challenge. In other words, group talk seemed to have become what Wegerif et al. (1998) call a clash between ideas, not between people.

4.1.5 Experimental adjustments

In this final paragraph we want to discuss certain aspects of the pilot study which have led to adjustments for the main study.

4.1.5.1 The lessons

As explained in Chapter 4, section 3.3, translated versions of the basic lessons developed by Mercer and Littleton (2007) were used. Despite adaptations of the teacher to make the content more appealing (while preserving the lesson goals), we found that some of the materials, like some of the stories, did not appeal to the pupils as might be expected. Because of this, we decided to adapt these for the main study, leaving learning goals untouched (cf. Appendix 1).

4.1.5.2 Lesson materials

Transcripts showed that the presence or absence of puzzles, writing sheets, booklets, etc during the exploratory phase of group work, i.e. the ten minutes we recorded, made a difference. For instance, when pupils had to discuss ten Children's Rights (three of which were made up by the teacher), their reasoning was much more visible than when they had to discuss a text and immediately write down answers to questions. Also, when the group received only one copy of learning material, pupils would move from their places when they had difficulty reading or seeing what the material was about. Especially in the first weeks, they would even pull materials out of each other's hands. This, too, had a negative effect on concentrated group talk. For the main study we therefore decided to develop a check list (cf. Appendix 4) which would enable the teacher to construct activities that would maximally induce reasoning by talk. Further, if learning materials were introduced, the teacher had to make sure each pupil had a copy.

4.1.5.3 Prompts for reasoning

The use of talk cards or 'prompts' during the basic lessons, which made some of the ground rules more tangible and which helped the pupils to adhere to the rules, was supposed to be replaced by a ground rule poster from week 4 onwards. However, transcripts of weeks 5–6 showed a drawback in the application of the ground rules. Moreover, some pupils specifically asked for the prompts, as if they felt the need for more scaffolding. Prompts were therefore reintroduced for two weeks and then again quietly omitted. This time the pupils did no longer ask for them. Therefore, in the main study we will maintain this flexibility and leave it to the teacher to determine whether or not pupils can use the prompts during group work after the basic lessons.

4.1.5.4 The learning process

Feedback and feed forward reflections on talk itself were part of each lesson. But although the teacher built in this kind of feedback during recorded lessons, it was not clear how and to what extent she did at other moments. Even more important is that whole class evaluation of talk tended to push pupils to give socially desired answers, which made giving proper feedback difficult and the feedback in itself less effective. For instance, when the teacher wanted to know who asked why-questions, nearly all pupils raised their hands in acknowledgment, while the transcripts show only a handful of them actually did. As giving feedback is known to have substantial learning effects (Hattie, 2009) it was obvious we needed it to be as uniformous as it could be in the main study. Therefore evaluation sheets were developed (see Appendix 1.4) for each triad but also for each individual. Using these sheets, pupils would fill in to what extent they had used the ground rules individually. After that, they would fill in a similar sheet at group level, comparing individual sheets. Only then would the teacher hear out the pupils and give them feedback. The use of these sheets would also give pupils a focus for the next group activity.

4.1.5.5 Raven's Progressive Matrices

After the pilot study we anticipated that repeating the pre- and post-tests would cause certain problems.

First, though the difference may be futile, giving pupils the uneven test first and the even test later (or vice versa) did not seem entirely fair. As Raven's gets more difficult with every item, we believe the chance existed that the uneven test is slightly easier than the even test. Second, and more important, after the pre- and post-tests it became clear that the pupils had scored quite high, as well at group level as at individual level. The maximum group scores, 17/18, did not leave much room for improvement. Knowing that the scores for Raven improve with age (Raven, 1998), it was conceivable that taking the same test from the - older - pupils in the main study might cause a ceiling effect. This effect would especially spoil the broth for Raven's Coloured Progressive Matrices, in which the more difficult C-, D- and E-levels are missing. In short, the expected ceiling effect would negatively intervene with the results of the tests in the subsequent experiments and was to be avoided. We did some inquiries in order to find a parallel version of Raven's Coloured Progressive Matrices but found none. In the end we decided to colour the C-, D- and E-levels of Raven's Standard Progressive Matrices ourselves using Photoshop Elements 2.0, added these to Raven's Coloured Progressive Matrices' A- and B-level and left out the AB-level. That way we had two tests of 60 items: a black and white one and a coloured one.

Given these alterations, tables 32 and 33 show the distribution of the Raven's test at pre- and post-test. In the tables 'x' indicates that a test item was provided, which means that of each set (A to E) pupils had to solve 6 items. Table 32 shows the distribution of test items for the pre- and post-test at group level. Likewise, table 33 shows test item distribution for individual pupils.

Table 32

Overview of the distribution of the test items at group level

Pre- and post-test at group level											
	'even' pre-test, black/white						**'uneven' post-test, black/white**				
	SET A	**SET B**	**SET C**	**SET D**	**SET E**		**SET A**	**SET B**	**SET C**	**SET D**	**SET E**
Item 1		x		x		Item 1	x		x		x
Item 2	x		x		x	Item 2		x		x	
Item 3	x		x		x	Item 3		x		x	
Item 4		x		x		Item 4	x		x		x
Item 5	x		x		x	Item 5		x		x	
Item 6		x		x		Item 6	x		x		x
Item 7		x		x		Item 7	x		x		x
Item 8	x		x		x	Item 8		x		x	
Item 9		x		x		Item 9	x		x		x
Item 10	x		x		x	Item 10		x		x	
Item 11	x		x		x	Item 11		x		x	
Item 12		x		x		Item 12	x		x		x

Table 33

Overview of the distribution of the test items at individual level

Pre- and post-test at individual level											
	'even' pre-test, coloured						**'uneven' post-test, coloured**				
	SET A	**SET B**	**SET C**	**SET D**	**SET E**		**SET A**	**SET B**	**SET C**	**SET D**	**SET E**
Item 1		x		x		Item 1	x		x		x
Item 2	x		x		x	Item 2		x		x	
Item 3	x		x		x	Item 3		x		x	
Item 4		x		x		Item 4	x		x		x
Item 5	x		x		x	Item 5		x		x	
Item 6		x		x		Item 6	x		x		x
Item 7		x		x		Item 7	x		x		x
Item 8	x		x		x	Item 8		x		x	
Item 9		x		x		Item 9	x		x		x
Item 10	x		x		x	Item 10		x		x	
Item 11	x		x		x	Item 11		x		x	
Item 12		x		x		Item 12	x		x		x

Evidently, all this included a form of re-testing, with a three-months interval. As re-testing was also and even more explicitly done in the original research we did not expect this to be a methodological flaw in the main study. However, one thing we had to take into account was this: earlier we decided that six triads would first be control groups and then become target groups (switching replication, cf. Chapter 4, section 1.2.2). This would mean that these pupils would do each of the Raven's tests three times, risking a strong re-testing effect, as the following table shows:

Table 34

Test organisation control and target groups, switching replication

Status	Week 1 = pre-test control group	Week 16 = post-test control group = pre-test target group	Week 30 = post-test target group
Triads	Raven uneven b/w	Raven even b/w	Raven uneven b/w
Individual pupils	Raven even coloured	Raven uneven coloured	Raven even coloured

We did not find a satisfying solution for this. It was either going for triple testing or risking the already mentioned ceiling effect, so we chose the former. Knowing this, we would have to look at the results of the target groups in week 30 with extra caution. E.g. as these pupils would do Raven's for the third time, we would expect them to perform better than the other pupils of the main study. We did consider splitting up Raven's in 20 items per test, but then the number of items would be so much reduced that it would be difficult to obtain significant results. Also, the validity of the test might be impaired, as is suggested in Raven's handbook (Raven, 1998), i.e. the items are ordened in such a manner that pupils learn by each item, gradually developing the thinking skills they need to solve the increasingly difficult items. We believe that leaving out two items out of three (resulting in 20 items) would be more liable to affect this continuity than when one out of two was left out (resulting in 30 items).

Finally, if a learning or re-testing effect is to remain completely absent, literature advises a 'safety zone' of two years (Schittekatte, 2000). Having to wait this long would, of course, be unpractical. Based on its own research, Raven found that test/re-test reliability is at its hightest (,88) after one week and at its lowest (,55) after one year. Re-testing after a year, let alone two, might thus induce less reliable results. Therefore, Raven's considers a three month interval for re-testing as reasonable. Furthermore, it should be noted that, like in the original study, we are not using Raven's to do entirely what it was constructed for, i.e. the measurement of IQ, but to measure pupils' problem solving skills.

4.1.5.6 A non-curricular discussion

In order to make pupils discuss a topic as intensely as possible, we wanted it to be familiar, i.e. close to their experience and to their foreknowledge. For the pre-test we therefore gave them three questions to tackle, followed by a prompt for discussion: a) How would you make the playground attractive to every pupil in your school? b) Which of your ideas would be the easiest / the hardest to accomplish and why is that? and c) Make a top three of your most feasible ideas.

Observations show that all pupils were immediately very much involved, which lead to lively discussions. However, transcripts show that the first questions stimulated most triads to employ only cumulative talk, as they started a mere brainstorm of possible ideas. Only as soon as they started to answer b) real discussions started. In the post-test we gave the pupils propositions which gave rise to discussion much sooner. Consequently, we must consider the results of pre- and post-test regarding the non-curricular discussion with some caution, as the increase of exploratory characteristics may also have been influenced by the test construct. For this reason, we decided to change the test construct and give pupils of the main study in both pre- and post-test only propositions they had to really argue about (see Chapter 4, 3.4.2).

4.1.5.7 Key words and the organisation of ideas

In order to connect ideas in talk, e.g. a claim with one or more arguments or with an argument and a conclusion, we use connectives like 'en', 'omdat', 'daarom', 'ook' ('and', 'because', 'therefore', 'also'). When they are used in the proper context, we call these in-context key words or key words in context, meaning key words for exploratory talk. The results of the non-curricular discussion show a strong increase of these key words after the intervention. In the problem solving task key word use is less apparent, because pupils use more implicit and semi-implicit arguments, as explained above. This does not mean, however, that talk is less exploratory. Evidence of this was given by Herrlitz-Biro et al. (2013), who identified exploratory talk in sequences almost void of key words. They concluded that key word count is in itself insufficient to fully grasp the 'exploratory nature' of a conversation, it takes strong qualitative analyses to complement. We believe that categorising arguments on several levels as in Rojas-Drummond and Zapata (2004), as we have decided to do, is such an analysis. As the number of arguments is also considered an indicator of exploratory talk, it may be interesting to combine both in a quantitative approach as well, which is what we will do in the main study.

The qualitative analyses of the pilot study has also yielded key words which we did not anticipate upon. Some are mere synonyms, other are paraphrases. These will we included in the NVivo 10 queries used for counting key words (cf. Appendix 5).

4.1.5.8 Individual variables

Based on this pilot study we have decided to include GOK- (or SES-)indication as a variable in order to answer RQ 5. The results of PISA studies (OECD, 2016) show that the gap between highest and lowest performing Flemish pupils is of the largest of all OECD countries and that GOK is one of the key elements in this. Taking GOK indicators into account may be helpful to clarify the role of speaking and listening skills as regards learning.

4.2 Main study

In this section we will address the results of the main study (cf. figure 14). First, background information is provided about the participating schools, teachers and pupils (4.2.1). Next, tests results are discussed concerning the use of exploratory talk (4.2.2) and the pupils' problem solving skills (4.2.3). The impact of individual variables is discussed in 4.2.4. A summarising paragraph (4.2.5) concludes this section.

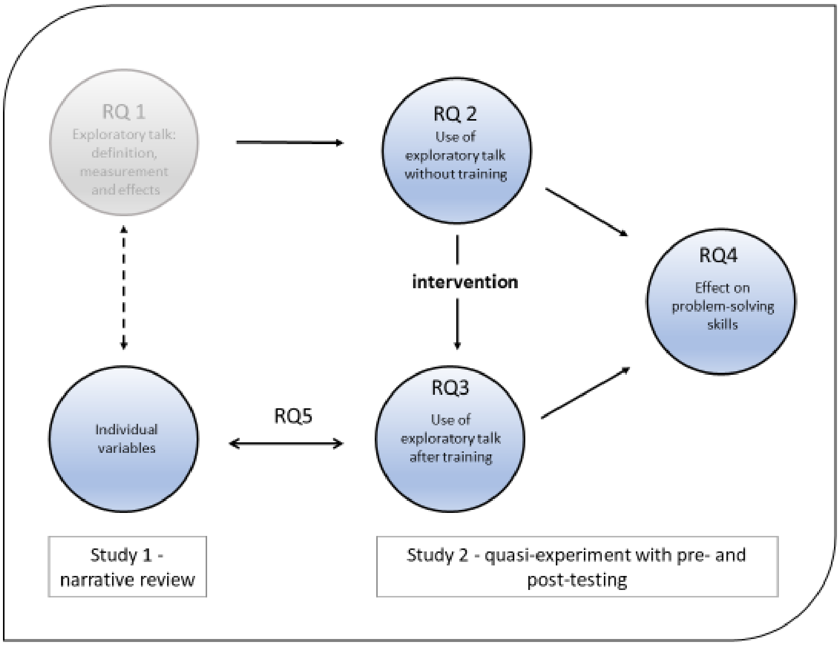


Figure 14. Global Research Design: focus on RQ 2-5

4.2.1 Background information

Schools 2-5 are local primary schools in districts around the city of Antwerp. Most pupils are indigeneous, the main differences between them being socioeconomic. In each school a small number of pupils has weaker language skills, learning problems or behavioural problems.

Except one, all teachers are female and their teaching experience varies from 2 to 22 years. Variation is also visible in their didactic approach but every teacher has at least basic experience with collaborative activities. Most teachers organise group work on a very regular basis, sometimes combining it with independent work and learning. None have any experience with exploratory talk nor have they heard of it in a pedagogic context.

Control and target classes are comparable as far as gender is concerned and so is the (small) number of pupils with GOK indicators. Pupils with weaker language skills, learning problems or behavioural problems have not been assigned to triads that were selected to be recorded.

4.2.1.1 The schools

School 2: 2 control groups = 2 target groups

Primary school 2 is a local school in a suburb of Antwerp. The school counts 236 pupils. Combined with the main school it is attached to there are 830 pupils in kindergarten and primary school, totalling 41 classes (16 in kindergarten and 25 in primary school) which are taken care of by 50 classroom teachers, one P.E.-teacher, one policy support assistant, one special needs coordinator, a secretary and a school director.

The 236 pupils of the local school nearly all live in the school's immediate neighbourhood. Except for a small minority they are Flemish/Belgian. Five pupils of the target group are Dutch, one pupil is Spanish and eight pupils are Belgian but have non-European parents. Most pupils are from middle-class families and all pupils speak Dutch fluently.

The school's pedagogy has been inspired by the catholic pedagogy for primary education of the 'Broeders van Liefde' and comprises five main goals:

1. Care for quality education. Every child must be given the opportunity to develop its talents. In order to realise this, the school must provide maximal professional guidance for education and cognitive development.

2. Respect. Every child, no matter its origin or socioeconomic status, is accepted and respected.

3. Christian meaning giving and perception. Pupils are expected to attend prayer service and meetings, and other sacramental celebrations organised by the school. The school respects every pupil's religion without sacrificing its catholic identity.

4. Professionalism. The school strives for dynamic, professional development and facilitates innovative thinking. Teachers are given ample opportunities to grow as professionalists via e.g. in-service training.

5. A close community. The school presents itself as a close community, fully integrated in the local parish. Maximal cooperation is strived for between pupils, teachers, parents and the school's directorate.

In the school year 2014-2015, i.e. the time when the experiment took place, the school focused on two specific projects: media and language policy.

School 3: 1 control group, 1 target group

School 3 is a small local school (233 pupils) in a suburb of Antwerp. All pupils were born in Belgium, have the Belgian nationality and speak Dutch as their mother tongue. 11,5% of the pupils' mothers speak a different mother tongue, however.

The school's pedagogy is Lasallian, which is based on a Christian view on education with special attention for personal development and participation. The school wants to be a warm home for every pupil. The teachers are invited to engage in contemporary, high quality education with an open eye for an ever changing world.

School 4: 1 control group, 1 target group

School 4 is a more rural school which lies between the cities of Antwerp and Sint-Niklaas. 308 pupils are divided over 18 classes which are taken care of by 28 teachers. Most pupils are middle-class children nearly all of whom are Flemish/Belgian (the total number of non-Belgian pupils is less than 5%).

The school shares its pedagogical project with the school 5. Within this project pupils are stimulated to become harmonically developed adults who share a positive self identity and who are motivated to become active citizens in service of the community they live in.

The school wants to offer high quality education which has the following characteristics: individual pupil guidance, systematic attention for holistic development, strong focus on motivation, initiative and learning-to-learn, and sound assessment via valid testing. Cognitive development goes hand in hand with the development of social, motoric and emotional skills. Teachers make sure that all aspects of learning are integrated in daily classroom practice and that each pupil is given opportunities to develop his own talents. In order to realise this powerful learning environments are created, while the teachers take up a responsible coaching and scaffolding role. The school wants to prepare its pupils for adult life as good as possible. In the end, pupils must be able to take full responsibility for their actions, to function as independent individuals and to develop a flexible attitude towards themselves and the people they interact with. To that end:

- pupils are encouraged and given positive stimuli
- pupils are encouraged to share lesson materials and interact in a respectful way
- teachers present themselves as role models
- teachers are all ears when pupils need them
- process assessment is equally important as product assessment
- conflicts are considered as opportunities for learning social skills
- the school has worked out a detailed health policy
- the school has made a priority of parent participation

In school year 2014-2015 the school focused on two specific projects: the realisation of horizontal and vertical curricula for world orientation and talent development.

School 5: 1 control group, 1 target group

School 5 is a local basisschool in a suburb of Antwerpen (left bank of the Schelde), close to school 5. The school counts 380 pupils (129 in Kindergarten, 251 in primary education). The primary school children are divided in 12 classes which are taken care of by 13 teachers. There is also one special needs coordinator who is assisted by 3 special needs teachers. Most pupils are Flemish/Belgian (less than 5% of all pupil are non-European, but their number is rising rapidly), who represent a wide social mixture.

For the pedagogical project of the school, see the description of school 4. In the school year 2014-2015 the school focused on especially on mathematics.

4.2.1.2 The teachers

Teacher 2 – school 2 – control group 2 = target group 4

Teacher 2 is 23 years old and has been teaching in school 2 since 2013. At first, she taught the 3rd form children, but since September 2014 she has been teaching a 5th form class. Before being admitted to school 2 she taught in special education for half a year. Teacher 2 received teacher education at the high school of KHK Vorselaar (now Thomas More Kempen - division Vorselaar). She also had in-service training on the integration of social media in classroom activities. This school year, she will enlist in in-service training on how to approach muslim children during religion and arts classes. Outside school she is also a dance teacher.

Concerning didactic approaches teacher 2 is very much in favour of all sorts of collaborative activities, which she integrates in classroom activities nearly every day. Cooperative activities, in which pupils really need one another in order to get work done, are very important to her. She alternates these with whole-class teaching and with individual excercises, not in the least because she wants every pupil to have his share of scaffolding. Differentiation is also one of teacher 2’s favourites, which she realises by grouping pupils according to their skills for various school subjects. Strong groups are given direct instructions to work independently, weaker groups receive extra support. Stronger pupils are also given opportunities to help weaker pupils. Despite her already considerable experience with collaborative learning, teacher 2 is not familiar with the notion or pedagogy of exploratory talk.

Teacher 3 – school 2 – control group 3 = target group 5

Teacher 3 is 42 years old and has been teaching in school 2 for 19 years. He received high school teacher education and enlarged his professional know how via in-service training (e.g. cooperative learning in multicultural groups (CLIM), reading performance and various subject oriented courses).

Teacher 3's favourite didactic approach is whole-class teaching. Collaborative activities are integrated 'whenever I think it has an added value'. Teacher 3 is specialised in the implementation of the ADI-model, i.e. systematically structured short and direct instructions which set pupils to work as quickly and efficiently as possible. Differentiation is incorporated whenever the need arises. Teacher 3 favours project work, in which learning goals, activities and materials of various subjects are combined. In his school, teacher 3 is the coordinator of MOS (Milieuzorg Op School), a project which focuses on environmental awareness in all classes.

According to teacher 3, group work is not always successful in his class. The teacher used to work with CLIM[21], but, as none of his colleagues followed his example, he stopped doing so. For two years now, school 2 is involved in a project called 'Wonen op het dak', which focuses on multiculturalism and which comprises a lot of collaborative activities. Despite his long experience, the teacher had never heard of exploratory talk before.

Teacher 4/5 – school 3 – control group

Teacher 4 is 43 and shares a class with teacher 5, who is 46 years old. Both teachers have about 20 years of teaching experience in school 3 and share two educational goals: stimulating pupils to become independent learners and to collaborate in an atmosphere of mutual respect. The teachers have had in-service training about pupil participation at school level, and about cooperative activities at classroom level. Collaborative activities are multifold in their class. Right now, the school is working out a curriculum on collaborative learning, a project to which both teachers contribute to substantially. Both teacher 4 and 5 also support and coach their younger colleagues. The teachers have no experience with or knowledge of exploratory talk, though they emphasised applying some of the ground rules implicitly.

---

[21] The CLIM method - or tool - answers to the need of teachers for strategies that enhance language skills of non-Dutch speaking pupils. The tool focuses on participative equity, mutual responsibility and group ownership, but also on constructive content interaction. Feedback and feed forward are essential elements of this method, as they make pupils improve both cognitive and social skills as they reflect on their activities and those of their peers. Therefore, for its educational potential to be fully realised, it is recommended that pupils do CLIM activities during consecutive school years (Paelman, 2004).

Teacher 6 – school 3 – target group

Teacher 6 is 26 years old and has been teaching in school 3 for 3 years. She received high school teacher education at Karel de Grote Hogeschool in Antwerp. Teacher 6 finds it very important to maintain a highly structured approach during classroom activities. Her experience with collaborative activities is intermediate, as she finds these hard to fit into the structured learning paths she has outlined. She is very good at organising whole-class and individual activities and at the same time finds it very important that pupils feel comfortable at school. For this purpose she regularly chooses forms of dialogic teaching and practises Philosophy for Children, but exploratory talk is new to her.

Teacher 7 – school 4 – control group 8, target group 9

Teacher 7 (26) has been teaching in school 4 since 2011. She received her teacher training at the Katholieke Hogeschool Sint-Lieven in Sint-Niklaas. Since then, she has had several in-service trainings about various pedagogical topics. Teacher 7's didactic approach is quite varied: she mixes whole-class instruction, individual and partner work with group work and forms of dialogic teaching, and also includes personalised work forms, e.g. 'Happy hour', which comes down to 'contract work' in which pupils receive a mixture of personalised assignments, some of which are obligatory and others are optional, and some of which they have to do individually while others receive help of a peer or of the teacher.

Teacher 7 organises group work regularly, in which she sometimes integrates the CLIM method (see the description of teacher 3). She is unfamiliar with exploratory talk and its didactic approach.

Teacher 8/9 – school 5 – control group 7, target group 10

Teacher 8 (32) and teacher 9 (46) both teach a 5th form in school 5.

Teacher 8 has 10 years of teaching experience which she practised in four different primary schools. She received her teacher training at Katholieke Hogeschool Sint-Lieven in Sint-Niklaas. To this, she and added a BanaBa-certificate for special education at the Arteveldehogeschool in Gent and a certificate for a short-track transitional year on Pedagogy at KULAK, university of Kortrijk. She had in-service training on interventional teaching and Dutch as a second language.

Teacher 9 has been teaching in several primary schools for 22 years, including substantial experience as GOK-coordinator, and received teacher education at H. Pius X-instituut in Antwerp. It is her 7th year in school 5.

Both teachers share a number of teaching strategies and preferences. They both want to simulate pupils' motivation for learning, cooperative learning and differentiation via various teaching methods. They have no knowledge of exploratory talk.

4.2.1.3 The pupils

Table 35 shows the total number of pupils in the control and target group, and gives an overview of the independent variables which were collected in order to answer RQ 5.

Table 35
Number of pupils in the control and target group, independent variables

	Number	Age (in year)	Boys (n)	Girls	GOK	High level spelling	High level mathematics
Control group	95	10.54	55	40	10	35	40
Target group	83	9.92	45	38	10	38	44
Total	178	10.23	90	78	20	73	84

Table 35 shows that the total number of pupils participating in the main study is 178[22]. 95 pupils are part of the control group (55 boys, 40 girls), while the target group counts 83 pupils (45 boys, 38 girls). The average age of the control group is 10,54 years, that of the target group is 9,92 years, which leaves an approximate age difference between both groups of 6 months. In the control group 35 out of 95 pupils have a high score for spelling, 60 pupils have a lower score. For mathematics, 40 pupils of the control group and 44 pupils of the target group have a high score. 55 pupils of the control group and 39 pupils of the target group have a lower score for mathematics. In both control and target groups 10 pupils receive GOK support.

We will now discuss this in more detail. As in two classes switching replication was applied, we will discuss classes by number instead of grouping them according to their status in the experiment (control vs. target group). All data were collected two weeks before organising the pre-test.

Control group 2 = target group 4 (switching replication)

Table 36
Pupils control group 2 = target group 4

Pupil control group 2	Triad	Pupil target group 4	Triad	Age (y)	Gender	GOK (1 = yes)	LVS spelling*	LVS maths*
21	10	56	23	11	F	0	2	1
22	11	57	24	10	M	0	2	2
23	11	58	24	10	M	0	1	1
24	10	59	23	10	M	0	2	1

[22] As in two classes pupils (n = 70) were subject to switching replication (cf. Chapter 4. Section 3.2) the total number of pupils must be must be reduced by 35 in order to obtain the number of unique pupils (n = 143).

25	11	60	24	10	F	0	2	2
26	9	61	22	10	F	0	1	1
27	7	62	20	10	F	1	1	1
28	8	63	21	10	F	1	2	2
29	10	64	23	10	F	0	2	1
30	7	65	20	10	M	0	2	1
31	11	66	24	9	M	1	1	1
32	10	67	23	10	M	0	2	1
33	9	68	22	10	M	0	2	1
34	8	69	21	9	M	0	2	1
35	9	70	22	10	M	0	1	2
36	8	71	21	10	M	0	1	1
37	7	72	20	10	F	0	1	1

**1 means 'high score', 2 means 'lower score'*

Class 2/4 is a 5th form which counts 18 pupils, i.e. 10 boys and 8 girls. 4 pupils live in a family which has recent, non-European roots. 4 pupils are have already repeated a year. 3 pupils receive GOK-support, no pupils have special needs. 8 pupils have high spelling skills, 10 pupils have weaker spelling skills. For mathematics 14 pupils have high skills, while 4 have weaker skills.

Control group 3 = target group 5 (switching replication)

Table 37
Pupils control group 3 = target group 5

Pupil control group 3	**Triad**	**Pupil target group 5**	**Triad**	**Age**	**Gender**	**GOK**	**LVS spelling ***	**LVS maths***
38	13	73	26	10	M	1	2	1
39	17	74	30	10	M	0	1	1
40	14	75	27	10	F	0	1	1
41	16	76	29	10	F	0	2	1
42	14	77	27	10	F	0	2	2
43	13	78	26	10	F	0	2	2
44	16	79	29	10	M	1	1	1
45	12	80	25	10	M	0	2	2
46	17	81	30	10	M	1	2	2
47	15	82	28	9	M	0	1	1
48	14	83	27	10	M	0	2	1
49	13	84	26	10	M	0	1	1
50	12	85	25	9	F	0	2	1

51	17	86	30	11	M	0	1	2
52	12	87	25	10	F	0	2	2
53	15	88	28	10	F	0	2	2
54	16	89	29	10	F	0	1	2
55	15	90	28	9	M	0	1	1

**1 means 'high score', 2 means 'lower score'*

Class 3/5 is a 5th form which counts 17 pupils, i.e. 10 boys and 8 girls. 4 pupils live in a family which has recent, non-European roots. 3 pupils have already repeated a year, while 1 pupil is a year ahead of the rest of his class. 3 pupils receive GOK-support. No pupils have special needs.

Target group 6

Table 38
Pupils target group 6

Pupil	Class	Triad	Age	Gender	GOK	LVS Spelling*	LVS Mathematics*
91	6	34	11	M	0	2	2
92	6	37	11	M	0	1	2
93	6	34	11	M	0	2	2
94	6	33	11	F	0	1	2
95	6	33	12	M	0	1	2
96	6	37	11	F	0	2	2
97	6	37	11	M	0	1	1
98	6	31	11	F	0	2	2
99	6	33	11	M	0	2	2
100	6	38	11	F	0	2	2
101	6	36	11	F	0	2	2
102	6	38	11	F	0	2	2
103	6	32	11	F	0	2	2
104	6	35	12	F	0	1	2
105	6	35	11	M	0	2	2
106	6	34	11	M	0	2	2
107	6	32	11	F	0	2	2
108	6	35	11	M	0	1	2
109	6	38	11	M	0	2	2
110	6	36	12	F	0	2	2
111	6	31	12	M	0	2	1
112	6	32	11	M	0	1	1
113	6	31	11	M	0	2	2
114	6	36	11	F	0	2	1

**1 means ‘high score’, 2 means ‘lower score’*

Class 6 is a 6th form which counts 24 pupils, i.e. 13 boys and 11 girls. All pupils speak Dutch as their mother tonghue, though one pupil’s mother does not. 3 pupils are entitled to GOK-support. 4 pupils have repeated one school year. For spelling 14 pupils have high skills, 10 have weaker skills. For mathematics 16 pupils have high skills, 8 have weaker skills.

Control group 7

Table 39
Pupils control group 7

Pupil	Class	Triad	Age	Gender	GOK	LVS Spelling*	LVS Mathematics*
115	7	39	10	M	1	2	2
116	7	44	11	M	0	2	1
117	7	39	10	F	0	2	2
118	7	43	10	M	0	2	2
119	7	43	9	F	0	2	2
120	7	40	10	M	0	2	1
121	7	41	10	M	0	2	2
122	7	43	10	M	0	1	1
123	7	44	10	F	0	2	1
124	7	40	10	M	0	2	1
125	7	41	10	F	1	2	2
126	7	42	10	M	0	1	2
127	7	40	10	M	0	1	2
128	7	42	11	M	0	1	1
129	7	44	12	M	0	1	1
130	7	42	11	F	0	2	2
131	7	39	10	M	0	1	1
132	7	41	10	M	0	1	1
133	7	41	11	M	0	2	1

**1 means ‘high score’, 2 means ‘lower score’*

Class 7 is a 5th form which counts 19 pupils, 14 boys and 5 girls. Most pupils are Flemish/Belgian, but two pupils have non-European parents who do not speak Dutch at home. One pupil has repeated a school year. Two pupils are one school year ahead of their peers. Two pupils receive GOK support. For spelling skills 7 pupils have high scores, while the majority of this class (12 pupils) has weaker skills. 10 pupils have high mathematics skills while 9 pupils have weaker skills.

Control group 8

Table 40
Pupils control group 8

Pupil	Class	Triad	Age	Gender	GOK	LVS Spelling*	LVS Mathematics*
134	8	47	11	M	0	1	2
135	8	46	11	F	0	1	2
136	8	47	11	M	0	2	1
137	8	48	11	F	0	2	2
138	8	48	11	F	0	2	2
139	8	49	12	F	0	2	2
140	8	45	11	F	0	2	2
141	8	47	12	F	0	2	2
142	8	48	12	F	1	2	2
143	8	45	11	M	0	1	1
144	8	46	11	M	0	1	1
145	8	49	12	F	1	2	2
146	8	49	12	M	0	1	2
147	8	46	11	M	0	2	2
148	8	49	11	M	0	2	2
149	8	45	10	F	0	1	2
150	8	48	11	M	0	2	2

**1 means 'high score', 2 means 'lower score'*

Class 8 is a 6th form which counts 17 pupils who all have the Belgian nationality. The parents of two pupils are non-European, one of whom speaks French at home while the others speak Dutch. Two pupils receive GOK support. 3 pupils have repeated one school year, while another pupil is a year ahead of his peers. One pupil has difficulties functioning in collaborative activities because of a neurological deficiency. On average, spelling and mathematic scores are weak in this class: for spelling 6 pupils are highly skilled, for mathematics only 3.

Target group 9

Table 41
Pupils target group 9

Pupil	Class	Triad	Age	Gender	GOK	LVS Spelling*	LVS Mathematics*
151	9	52	9	F	0	2	2
152	9	50	10	M	0	2	2
153	9	52	10	F	1	2	?
154	9	54	9	F	0	1	1

155	9	51	9	F	0	1	2
156	9	53	11	F	1	2	2
157	9	50	11	M	0	2	1
158	9	54	11	M	1	2	2
159	9	53	10	M	0	2	2
160	9	52	9	F	0	?	?
161	9	51	10	F	0	1	1
162	9	53	10	M	1	?	2
163	9	51	10	M	0	2	1
164	9	54	10	M	0	2	2
165	9	50	10	F	0	1	1

**1 means 'high score', 2 means 'lower score'*

Class 9 is a 5th form which counts 15 pupils, 7 boys and 8 girls. 3 pupils have repeated one school year. 4 pupils receive GOK support. 4 pupils have high spelling skills, 5 have high mathematics skills. This means that the majority of the pupils has weak spelling skills (n = 9) and weak mathematics skills (n = 8). The spelling scores of two pupils for spelling and one for mathematics were not available.

Target group 10

Table 42

Pupils target group 10

Pupil	Class	Triad	Age	Gender	GOK	LVS Spelling*	LVS Mathematics*
166	10	58	10	F	0	1	2
167	10	57	10	M	0	2	2
168	10	59	9	M	0	1	1
169	10	60	10	F	0	2	2
170	10	60	9	M	0	1	2
171	10	56	10	F	0	1	2
172	10	57	10	F	0	1	1
173	10	58	9	M	0	1	1
174	10	59	10	F	0	2	1
175	10	59	11	M	0	1	2
176	10	56	10	M	0	1	1
177	10	60	10	F	0	1	1
178	10	56	10	M	0	2	1
179	10	58	11	F	0	2	2
180	10	57	10	M	0	2	1

**1 means 'high score', 2 means 'lower score'*

Class 10 is a 5th form which counts 15 pupils, 8 of which are boys and 7 are girls. All pupils are Flemish/Belgian. Two pupils have repeated a school year. No pupils receive GOK support. 9 pupils show high skills and 6 show weaker skills for spelling. For mathematics 7 pupils have high skills and 8 have weaker skills.

Control group 11

Table 43
Pupils control group 11

Pupil	Class	Triad	Age	Gender	GOK	LVS Spelling*	LVS Mathematics*
181	11	63	10	M	0	1	2
182	11	61	10	F	0	2	2
183	11	64	9	M	0	1	1
184	11	61	10	M	0	1	2
185	11	65	10	M	0	1	1
186	11	66	11	F	0	1	2
187	11	62	10	F	0	1	1
188	11	63	11	M	0	2	2
189	11	66	10	M	0	1	1
190	11	63	10	F	0	2	2
191	11	62	10	M	0	2	1
192	11	61	9	M	0	1	1
193	11	62	10	F	0	2	2
194	11	64	9	F	0	2	2
195	11	65	9	M	0	2	2
196	11	66	10	F	0	1	1
197	11	65	10	M	0	2	1
198	11	64	10	F	0	1	2

**1 means 'high score', 2 means 'lower score'*

Class 11 is a 6th form which counts 18 pupils, i.e. 10 boys and 8 girls. All pupils are Flemish/Belgian and speak Dutch for their mother tongue. 2 pupils have repeated one school year. No pupils receive GOK support. 10 pupils have high spelling skills, whereas 8 pupils have weaker spelling skills. Mathematics skills shows the reverse: 8 pupils have high skills, 10 pupils have weaker skills.

4.2.2 Test results: exploratory talk (RQ 2 and RQ 3)

In this paragraph we will answer RQ 2 and 3 by discussing the results of the pre- and post-test regarding the use of key words (4.2.2.1), turn-taking (4.2.4.2), long utterances (4.2.2.3) and the use and quality of arguments (4.2.2.4). But as we did earlier when discussing the pilot study, we first wish to present an illustrative analysis of two transcripts (with translations from Dutch).

Transcript 3[23]

Pre-test: problem solving activity (Raven's Progressive Matrices)

		English translation
Kim:	Laat dan maar, hé. Dat vierkantje *(wijst)* pak ik.	Leave it. I will take *(points)* that square.
Teresa:	*(wijst)* Nee, dat.	*(points)* No, this one.
Kim:	*(wijst)* Het is dit of dit.	*(points)* It's this or this.
Teresa:	*(wijst)* Ik denk dit.	*(points)* I think it's this.
Jef:	*(wijst)* Nee dat.	*(points)* No, that one
Teresa:	*(wijst)* Nee, hé. Het is dit.	*(points)* No, it's this.
Jef:	(leunt achterover op zijn stoel) Oké, Kim.	(leaning backwards on his chair) OK, Kim
Teresa:	We nemen zes.	We take six.
Jef:	Allé Kim. We moeten wel overeenkomen, hé.	Oh come one, Kim, we have to agree, don't we.
Teresa:	Het is zes, echt waar.	It's six, really.
Kim:	Oké, laat maar zitten. Als het fout is dan is het jullie fout, hé.	OK, leave it. But if it's wrong, then it's your fault.
Jef:	Gij zit in onze groep, hé Kim.	You are in our group, Kim.
Kim:	Maar ik had het meestal juist.	But I am usually right.
Jef:	Meestal. En ikke ook. In rekenen. Dat is een soort van rekenen.	Usually. And so am I. In mathematics. This is a kind of mathematics.
Kim:	(wijst de volgende opgave aan) Dit, drie.	(pointing at the next puzzle) This, three.
Jef:	Da's direct drie.	That's three alright.

This extract shows a lot of characteristics of disputational talk. Pupils are hardly using any arguments, turns are very short, there are no 'what do you think?' questions, let alone why-questions, and no joint agreement is reached. From the start Kim believes he knows the correct answer. As soon as he experiences objections he turns away from the conversation. When Jef tries to draw him back in, Kim only warns him and Teresa that he is not responsible should the answer be wrong. His argument is all about maintaining self-identity (Polo et al., 2015), e.g. he says that he often has the answers right, probably in other contexts. Jef then reacts in the same way, saying he also has many answers right, especially in mathematics (with which he compares the puzzle). Teresa makes her point in the first half of the conversation and even tries to force a decision, but she withdraws from the conversation as soon as Kim and Jef start convincing one another how good they are. Eventually Kim initiates the next puzzle and

[23] All names are fictitious.

proposes a solution (to which Jef immediately agrees). Exploratory talk is mostly absent in this conversation.

The following transcript illustrates a post-test conversation.

Transcript 4[24]
Post-test: problem solving activity (Raven's Progressive Matrices)

		English translation
Teresa:	Ik denk twee. Kim?	I think it's two. Kim?
Kim:	Ik denk vijf.	I think it's five.
Teresa:	Waarom denk jij dat?	Why do you think that?
Jef:	Wacht nee…	Wait, no…
Kim:	Maar kijk… *(wijst)* Je hebt een patroon..	Look… *(points)* You have this pattern
Jef:	*(wijst)* Een rond vierkant, een vierkant met een kruis door, een vierkant met bolletjes. Zo eentje met rond, zo eentje met een kruis door, zo eentje met bolletjes. Dus dat moet deze dan hè. Hè want … Zo en er is nog geen met een rondje ervan hè?	*(points)* A round square, a square with a cross inside it, a square with small circles. Here's one round, one with a cross going through it, one with circles. So this should be, because … Like that and there is none with a circle, don't you think?
Kim:	Maar met…	But with …
Teresa:	Zie jij ergens … *(wijst)* zo kijk hè, wacht hè, doe uw vinger eens weg. Euhm… Zie jij ergens zo eentje?	Do you see anywhere… *(points)* like this, wait, take your finger from it. Do you see one like this?
Teresa:	Maar dan met een andere tekening?	You mean with a different drawing?
Jef:	Neen.	No, I don't
Kim:	Neen.	No, I don't
Teresa:	Neen, dus dat staat er niet bij.	No, so that is not part of it.
Jef:	*(wijst)* Dat zal deze zijn eigenlijk hè.	*(points)* In that case it must be this, don't you think?
Kim:	Ja, ja… .	Yes, yes
Teresa:	Zijn we het dan eens?	So, do we agree?
Jef:	Ja.	Yes.

This dialogue is much more exploratory than the first one. Pupils ask for each other's opinion, ask for arguments and counterchallenge claims. Overall, there is less pointing and more explicit conversation, turns are longer and the group reaches a joint agreement. Teresa initiates the conversation by expressing her solution but she immediately asks if Kim agrees. Kim then suggests another solution. Jef, who was very competitive and quickly bad-tempered in the pre-test still, reacts impulsively but now he does so by formulating an elaborate argument for his

[24] All names are fictitious.

claim. Teresa counterchallenges him, not so much by giving a counterargument but by questioning his logic. Then Jef understands and changes his claim himself, to which all agree.

Both fragments show how four language indicators can change the type of conversation from disputational to exploratory (and vice versa): key word use in context, turn-taking, the length of utterances and the use of arguments. In the first transcript key words like ‘why’, ‘because’, ‘so’, ‘agree’, etc are lacking and consequently, so are (elaborate) arguments and long utterances, while turn-taking shows an asymmetrical pattern (Teresa withdraws halfway). In the second transcript a why-question generates an elaborate argument (‘Look…’, ‘So….’, ‘Because…’, ‘Don’t you think?’) which is also a very good example of a long functional utterance. Additional questions (‘Do you see…?’, ‘You mean with …?’) refine this argument and eventually a conclusion is drawn (‘In that case…’) after which the group’s consent is asked (‘So, do we agree?’). All the time, all group members remain involved and turn-taking is more symmetrical.

In the next paragraphs we will discuss the global presence of these indicators in the pre- and post-test conversations of the control and target group.

4.2.2.1 Use of key words in context

In this paragraph we will first compare the use of key words for exploratory talk for the observed triads during the discussion of a non-curricular topic (10’), before and after the intervention. Next, similar results will be presented as the triads were solving the Raven’s Progressive Matrices, also during the pre- and post-test. Of each conversation ten minutes were analysed.

Using NVivo 10 we counted the number of key words in each conversation. The conversations of the control and target group triads were prepared for analysis of their non-curricular and problem solving discussion. Due to a technical flaw the non-curricular discussion in one triad of the target group could not be transcribed.

Absolute vs. proportional key word use

As mentioned earlier (see 3.5.1 c) an absolute count of key words does not suffice to draw conclusions about the use of exploratory talk. Pupils may very well use words like ‘I think…’, ‘would…’, ‘but…’, etc for purposes other than exploring a line of thought. Such non-exploratory contexts can be cumulative or disputational, or perhaps examples of Brevig’s (2006) *reflective* or *operational talk*, or even suggest others types of talk which may not have been defined yet. For that reason, in this study only key words used in exploratory context (in short: key words in context) should be counted. Thus, after all key words were identified, two independent researchers marked those key words used in context through qualitative analysis (Cohen’s kappa = 92[25]). Key words that were used in a non-exploratory context and so did not reflect the use of exploratory talk, were left out in a second count.

Before analysis we expected that some triads might already be using many key words in other contexts, whereas other triads might perhaps be using few key words but mostly in an exploratory context. Therefore, before turning to the counting results of key words in context (or in-context keywords), we want to show how many of all key words found in a conversation were retained after we left out all key words that were used out of the exploratory context (out-of-context keywords or keywords out of context).

Table 44 shows the absolute use of in-context key words vs. out-of-context key words, in percentages. The table should be read as an answer to the following question: if the total number of key words found in all conversations is 100% what percentage remains after ruling out those key words that are used out of context, both in the pre- and post-test?

Table 44

Absolute use of key words in context, pre- and post-test, non-curricular and problem solving discussion, control and target group (in %)

			Non-curricular discussion		**Problem solving discussion**	
Status			**Pre-test**	**Post-test**	**Pre-test**	**Post-test**
Control group	N	Valid	15	15	15	15
		Missing	0	0	0	0
	Mean		81.516	77.693	85.034	82.967
	Std. Deviation		8.769	6.689	6.172	7.466
	Range		28.02	24.41	19.43	24.04
	Minimum		64.79	64.55	72.88	71
	Maximum		92.81	88.96	92.31	95.04
Target group	N	Valid	14	15	15	15
		Missing	1	0	0	0
	Mean		85.639	86.854	83.33	92.509
	Std. Deviation		5.326	7.982	6.458	4.2605
	Range		18.49	32.86	20.81	13.87
	Minimum		74.26	61.9	71	83.54
	Maximum		92.75	94.76	91.81	97.41

[25] Cohen's kappa coefficient is a statistic which measures inter-rater agreement for qualitative (categorical) items. The kappa score can vary from 1 tot 0. Kappa assumes its theoretical maximum value of 1 only when both observers distribute codes the same. Very good agreement is indicated by a score between 1 and .80, good agreement by a score between .80 and .60, etc (Cohen, 1960).

Table 44 shows the following:

a) For the non curricular-discussion:

The average use of key words in context in the control group (n = 15) is about 81.5% of all key words found, which means that 18.5% are key words out of context. There are, however, huge differences between triads, as the standard deviation (s.d =8.77) and range (28) suggest. The lowest percentage of in-context key word use found in a triad is 64.8%, the highest is 92.8%. Apparently the first triad used only two thirds of its key words in an exploratory context. The second triad used more than nine out of ten of its key words in an exploratory context. Scores of the other triads lie between those two. This suggests that some triads use key words in context to a great extent, while others (also) use many key words out of context. This may imply that some triads have produced more exploratory talk from the start. This is important, because it may influence their margin for progress.

When we look at the results of the post-test, we see a decline in key word in context use to 77.6%. Standard deviation (s.d = 6.6) and range (24) are somewhat less, but the difference between minimal use (64.5%) and maximal use (88.9%) remains high.

The average use of key words in context in the target group (pre-test: n = 14; post-test: n = 15) during the non-curricular discussion is 85.6% which means that 14.4% are key words out of context. Again there are considerable differences between triads, though somewhat less than in the control group, as standard deviation (s.d = 5.32) and range (18.5) suggest. The lowest percentage of in-context key word use found in a triad is 74.2%, the highest is 92.8%. Again this may suggest that some triads have more 'exploratory work' to do than others.

After the intervention, there is a slight increase of relative key word use in context (M = 86.8) and a strong increase of standard deviation (s.d = 7.9) and range (32.8). This may suggest that some triads have picked up the trail of exploratory talk while others have not or less. Some even seem to have gone off the exploratory track. This is reflected in the maximum and minimum scores for the use of key words in context, which now range from 94.7% to 61.9%.

b) For the problem solving discussion:

The average use of key words in context in the control group (n = 15) is about 85%, which means that 15% are key words out of context. Again, triads show a lot of variety in their use of key words in context, as the standard deviation (s.d = 6.2) and range (19.4) suggest, though less than in the non-curricular discussion. The lowest percentage of in-context key word use found in a triad is 72.9%, the highest is 92.3%.

After the intervention, the use of key words in context decreases to 83%. Standard deviation (s.d = 7.4) and range (24) are comparable to the results of the pre-test, but the difference between minimal use (71%) and maximal use (95%) remains high.

The average use of key words in context in the target group (n = 15) is 83.3%. This means that 16.7% are key words out of context. The differences between triads are again obvious: standard deviation (s.d = 6.4) and range (20.8) are high. The lowest percentage of in-context key word use found in a triad is 71%, the highest is 91.8%.

Here the intervention shows clear effects, however: there is an increase of relative key word use in context (M = 92.5) and a decrease of standard deviation (s.d = 4.2) and range (13.9). Moreover, minimal usage has gone up from 71 to 83.5% and maximum usage reaches 97.4%. This suggests a collective evolution: on average, triads put key words in context to more use in the problem solving conversation after the intervention.

Significance

We performed a non-parametric Wilcoxon Signed Rank Test in order to find out whether the difference between key word use between pre- and post-test is significant in the target and control group (see table 45).

Table 45
Significance of increasing/decreasing percentage use of key words in context

Status		**Difference pre-/post-test non-curricular discussion**	**Difference pre-/post-test problem solving discussion**
Control group	Z	-1.099[b]	-1.136[b]
	Asymp. Sig. (2-tailed)	0.272	0.256
Target group	Z	-1.590[c]	-3.181[c]
	Asymp. Sig. (2-tailed)	0.112	0.001

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

As is shown in table 45, in the non-curricular discussion we found no significant evolution for either the control group (decrease with nearly 3.8%, $p = .272$) nor the target group (increase with 1.2%, $p = .112$). In the problem solving discussion, however, significance was found for the target group (increase with 9.2 %, $p = .001$) and not in the control group (decrease with 2 %, $p = .256$).

Table 46 shows how the scores of the control group and target group evolved.

Table 46

Evolution of the use of key words in context: control and target group

Status			Score difference non-curricular discussion	Score difference problem solving discussion
Control group	N	Valid	15	15
		Missing	15	15
	Mean		-3.823	-2.067
	Std. Deviation		11.290	6.948
	Minimum		-28.26	-13
	Maximum		14.62	11.36
Target group	N	Valid	15	15
		Missing	12	12
	Mean		6.924	9.179
	Std. Deviation		24.288	6.180
	Minimum		-17.75	-1.59
	Maximum		83.49	19.72

As the evolution of key words use in the control group is negative for both conversations (-3.8 and -2.1% respectively), while the target group shows clear progress (+6.9 and +9.18% respectively), we did a Mann Whitney U Test to determine the differences between both groups. The evolution fo the score difference (pre- vs. post-test) between both groups proved to be not significant for the non-curricular discussion ($p = .141$) but it is significant for the problem solving discussion ($p = .000$), as table 47 shows.

Table 47

Mann-Whitney U Test: significancy between control and targetgroup: evolution in the use of key words in context

	Score difference non-curricular discussion	Score difference problem solving discussion
Mann-Whitney U	77	24.5
Wilcoxon W	197	144.5
Z	-1.472	-3.65
Asymp. Sig. (2-tailed)	0.141	0.000
Exact Sig. [2*(1-tailed Sig.)]	.148[b]	.000[b]

a. Grouping Variable: TARGET GROUP (1) CONTROL GROUP (0)

b. Not corrected for ties.

Based solely on the increasing presence of the indicator key words in context this would mean that pupils of the target groups have learnt to talk significantly more exploratively during their problem solving conversation than the pupils of the control group.

Proportional use of key words

Further, as the number of words used in a conversation can vary considerably among triads, we felt it necessary to give an overview of the key word count in percentages that reflect the proportion or relative use of key words in each conversation. Proportional numbers also make it possible to compare the use of key words among triads in a more objective way. This is even more important when we combine these with the results for other indicators of exploratory talk, e.g. the number and quality of arguments.

Frequency table 48 shows the proportional use of key words in context as related to the length of the conversations (percentage of key words), also expressed in percentages. This means that, in this table, any number behind 'Mean' should be interpreted as the average percentage of key words in context used in all conversations.

Table 48

Proportional use of key words in context, pre- and post-test, non-curricular and problem solving discussion, control and target group

			Non-curricular discussion		Problem solving discussion	
Status			**Pre-test**	**Post-test**	**Pre-test**	**Post-test**
Control group	N	Valid	15	14	15	15
		Missing	15	16	15	15
	Mean		11.503	11.657	12.466	11.123
	Std. Deviation		3.259	3.448	2.661	2.331
	Range		10.81	14.76	9.85	7.8
	Minimum		6.1	5.51	7.93	7.17
	Maximum		16.91	20.27	17.78	14.97
Target group	N	Valid	14	15	15	15
		Missing	13	12	12	12
	Mean		13.238	13.217	11.087	13.686
	Std. Deviation		2.547	2.617	1.835	2.583
	Range		10.49	10.08	6.35	8.94
	Minimum		9.78	6.77	7.56	9.7
	Maximum		20.27	16.85	13.91	18.64

Table 48 shows the following:

a) For the non-curricular discussion:

The average use of key words in context in the control group (n = 15) comprises 11.5% of the conversations. Standard deviation is 3.26, range is 10.8. The minimal key word percentage is 6.1%, while the maximum is 16.9%. After the intervention, the results are somewhat similar. The use of key words in context very slightly increases to 11.6%. Standard deviation (s.d = 3.4) and range (14.7) increase: minimal use drops slightly to 5.5% while maximal use reaches a higher 20.3%.

The average use of key words in context in the target group (n = 14 during pre-test and 15 during post-test) is higher than in the control group, i.e. 13.2%, with a lower standard deviation (s.d = 2.5) and nearly similar range (10.49). Minimum (9.8%) and maximum (20.3%) use are also higher compared to the control group. After the intervention, scores are very similar to the pre-test ones: average use is 13.2%, standard deviation is slightly higher (s.d. = 2.6%), range is slightly lower (10%), while the minimum (6.8%) and maximum (16.85%) scores drop. The generalising suggestion that some triads have picked up exploratory talk after the intervention while others have not (see the comments on table 44) is difficult to uphold. In other words, when we consider proportional use, the target group does not show any evolution in key word use during the non-curricular discussion.

b) For the problem solving discussion:

The proportional use of key words in context in the control group (n = 15) is about 12.4%. Variety is reflected by a moderate standard deviation (s.d = 2.7) and range (9.85), which is less than in the non-curricular discussion. The lowest percentage of in-context key word use found in a triad is 7.9%, the highest is 17.8%. After the intervention, the use of key words in context decreases to 11.1%. Standard deviation (s.d = 2.3) and range (7.8) decrease as well. Minimal use remains practically the same as in the pre-test (7.2%), but maximum use drops to 15%.

The average proportional use of key words in context in the target group (n = 15) is 11.1%, which represents a lower start compared to the control group. The differences between triads are reasonable, as standard deviation (s.d = 1.8) and range (6.3) show. The lowest percentage of in-context key word use found in a triad is 7.6%, the highest is 13.9%. Here the intervention shows clear effect, as the relative use of key word in context increases (M = 13.7%). Standard deviation rises (s.d = 2.6) and so does range (8.9). Moreover, minimal usage has gone up from 7.6 to 9.7% and maximum usage reaches 18.6%, coming from 13.9%. The collective evolution suggested by table 44 is confirmed: on average, triads put key words in context to more use in the problem solving conversation after the intervention, but not all triads have made equal progress.

Significance

We performed a non-parametric Wilcoxon Signed Rank Test in order to find out whether the difference between key word use between pre- and post-test is significant in the target and control group (see table 49).

Table 49
Significance of proportional key words use in the pre-test vs. post-test

Status		**Difference pre-/post-test non-curricular discussion**	**Difference pre-/post-test problem solving discussion**
Control group	Z	-.220[b]	-1.647[b]
	Asymp. Sig. (2-tailed)	0.826	0.100
Target group	Z	-.408[c]	-2.897[c]
	Asymp. Sig. (2-tailed)	0.683	0.004

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

As table 49 shows, the only significant difference (p = .004) can be found in the target group and comprises proportional use of key words in context during the problem solving conversation. Other differences, both in the experiment and control group, are not significant. Based on a Mann-Whitney U Test we found significance between the experiment and control group for the problem solving conversations as well (p = .008), but not for the non-curricular discussion (p = .110).

Summary: use of key words in context

Absolute use of key words in context

Our hypothesis was that if pupils are trained in the use of exploratory talk their use of key words in context would increase and the number of key words out of context would decrease. This is confirmed partially by our data. The use of key words in context, i.e. exploratory talk, increases significantly and collectively in the problem solving discussions of the target group, while it decreases in the control group. This evolution is less outspoken when we examine the non-curricular discussions, where no significant increase was found in the target group (in the control group there is a non-significant decrease).

These outcomes may appear ambiguous and even contraditory, but we believe they can be explained. First, exploratory talk was trained mostly during problem solving tasks which are more in line with Raven's Progressive Matrices, the problem solving pre- and post-test. Second, although pupils used a considerable number of key words in context during the non-curricular discussion, a discussion in itself often lacks some of the principles of problem solving. For instance, in an open discussion coming to a consensus (ground rule 6) is less important than in a problem solving discussion. Also, an open discussion often includes characteristics of disputational talk. Qualitative analysis and comparison with the occurrence of other indicators of exploratory talk should cast more light on this issue.

Proportional use of key words in context

The proportional use of key words in context during problem solving discussions increases significantly after the intervention in the target group. Key word use during the non-curricular discussion does not show any positive evolution. The difference in key word use between control and target groups at the start of the experiment cannot be explained by our data. As the standard deviation among triads in the target group has decreased, it is reasonable to assume that the pupils of the target group have evolved more collectively, as we commented earlier. Difference in progress between triads remains huge.

4.2.2.2 Turn-taking

As pupils were mastering exploratory talk we anticipated they would interrupt one another less than in e.g. disputational talk and exchange turns in a more democratic way (cf. ground rule 1: 'What do you think?'). We also expected pupils to be less dominant or recessive during a conversation as group identity would become more important than self-identity (Polo et al., 2015) and as a result of conversation becoming more exploratory (Littleton et al., 2005; Wegerif et al., 2005). Therefore we counted turns in each triad and calculated how these turns were distributed within the groups, before and after the intervention.

As explained earlier (see section 3.5.1 in this chapter) we chose two arbitrary cut-off points: one to decide whether a conversation is a/symmetrical, another to determine interactive dominance/recession. This means that quantitative symmetry is established when the difference between the pupil taking the most turns and the pupil taking the least number of turns is not more than 12% of all turns. Interactive dominance happens when one pupil takes 45% or more of all turns, while interactive recession means that a pupil takes less than 20% of all turns. Mathematically, this means that every conversation in which interaction is dominant or recessive, is asymmetrical by definition, but not all asymmetrical conversations need to include interactive dominance/recession.

Quantitative symmetry

Table 50 shows the percentages of turn-taking per pupil and per triad during problem solving conversations (Raven's Progressive Matrices) in the pre- and the post-test. It also shows the number of asymmetrical discussions (+/- 12%) and the number of instances of interactive dominance (+45%) or recession (-20%).

Table 50
Number of conversations showing quantitative asymmetry (+12% spread)

			Non-curricular discussion: quantitative asymmetry		**Problem solving discussion: quantitative asymmetry**	
Status			**Pre-test**	**Post-test**	**Pre-test**	**Post-test**
Control group	Valid	0	9	6	7	7
		1	6	6	8	8
		Total	15	15	15	15
Target group	Valid	0	2	7	5	8
		1	12	8	10	7
		999	1			
		Total	15	15	15	15
100	Valid	0	3	3	3	3

Table 50 shows the following:

a) For the non-curricular discussion:

Before and after the intervention 6 triads of the control group (n = 15) show quantitative asymmetry. In the target group (n = 14 in the pre-test, n = 15 in the post-test) quantitative asymmetry affects 12 triads in the pre-test, decreasing to 8 triads in the post-test.

b) For the problem solving discussion:

Before and after the intervention 8 triads of the control group (n = 15) show quantitative asymmetry. In the target group (n = 15) quantitative asymmetry affects 10 triads in the pre-test, decreasing to 7 triads in the post-test.

Significance

As for quantitative symmetry we found near significant progress for the target group in the non-curricular discussion (Wilcoxon Signed Rank Test; $p = .059$) and non-significant progress in the problem solving discussion (Wilcoxon Signed Rank Test; $p = .257$).

Interactive dominance/recession

In the next table we examine interactive dominance/recession.

Table 51
Number of conversations featuring interactive dominance/recession

			Non-curricular discussion		Problem solving discussion	
Status			**Pre-test**	**Post-test**	**Pre-test**	**Post-test**
Control group		0	14	14	13	13
	Valid	1	1	1	2	2
		Total	15	15	15	15
Target group		0	12	12	8	14
	Valid	1	3	3	7	1
		Total	15	15	15	15
100	Valid	0	14	14	13	13

Table 51 shows the following:

a) For the non-curricular discussion:

Before and after the intervention interactive dominance/recession is visible in one triad of the control group (n = 15). The intervention brings no change to that. In the target group (n = 15) three triads show interactive/dominance recession before and after the intervention. Again the intervention seems to have made no difference.

b) For the problem solving discussion:

Before and after the intervention interactive dominance/recession appears in two triads of the control group (n = 15). After the intervention nothing has changed. In the target group (n = 15), however, zeven triads show interactive/dominance recession before the intervention but after the post-test only one triad is affected.

Significance

The only significant evolution we found was for the problem solving discussion, where the presence of interactive dominance/recession was almost reduced to zero in the target group (Wilcoxon Signed Rank Test; $p = .034$).

Summary: turn-taking

Quantitative symmetry

Quantitative asymmetry dominates the target groups during the non-curricular discussion in the pre-test, but after the intervention the number of triads showing quantitative asymmetry roughly equals the ones showing quantitative symmetry. Simply put, turn-taking has become more democratic for a number of triads. We do not see this evolution in the control groups, but here fewer groups start off asymmetrically during the pre-test. We do see a positive evolution in the problem solving discussion. Here, though change is moderate, the target group exhibits less quantitative asymmetry after the intervention than before, whereas the results of the control group remain the same.

Interactive dominance/recession

We expected conversations to become democratic (with less or even no interactive dominance/recession) after the intervention in the target group. This worked out significantly for the problem solving discussion in this group. The control group shows no evolution at all, neither were there any relevant interactive changes during the non-curricular discussions, though it must be said that interactive dominance/recession was rare from the beginning. The difference at the pre-test for the problem discussion between the control and target group is remarkable but cannot be explained by our data.

4.2.2.3 Length of utterances

Tables 52-55 show the number of long utterances. As discussed in Chapter 4, section 3.5.1, we decided to make calculations for three cut-off points which determine whether an utterance is long or not: >=70, 100 and 115 characters). Tables 52 and 54 show the pre- and post-test results for the non-curricular discussion, while tables 53 and 55 show the pre- and post-test results for the problem solving test. We will first discuss the absolute number of long utterances and then look at the proportional number.

Absolute number of long utterances

Table 52 describes the average frequency of long utterances in the non-curricular discussion, while table 53 zooms in on the problem solving discussion.

Table 52

Long utterances in the pre- and post-test of the non-curricular discussion

Status			**Pre-test 70 characters**	**Post-test 70 characters**	**Pre-test 100 characters**	**Post-test 100 characters**	**Pre-test 115 characters**	**Post-test 115 characters**
Control group	N	Valid	15	14	15	14	15	14
		Missing	0	1	0	1	0	1
	Mean		24	20.29	13.33	10.57	9.47	7,29
	Std. Dev.		12.61	8.722	9.424	4.783	7.492	3,891
	Minimum		9	9	1	3	0	2
	Maximum		52	39	31	22	24	17
Target group	N	Valid	14	15	14	15	14	15
		Missing	1	0	1	0	1	0
	Mean		21.86	25.6	12.29	16.53	9.21	13,07
	Std. Dev.		9.281	8.236	6.922	6.479	5.409	5,106
	Minimum		9	15	3	8	2	4
	Maximum		47	40	31	31	22	21

Table 52 shows that, for the control group, the absolute number of long utterances decreases in the post-test. It makes no difference if we define 'long utterances' as more than 70, 100 or 115 characters. For 70 characters: from 24 tot 20.29 utterances f; for 100 characters: rom 13.33 to 10.57 utterances; for 115 characters: from 9.47 to 7.29 utterances. Standard deviations are high for all scores and so is the range between minima and maxima.

In the target group we see the contrary: increasing numbers for the three cut-off points. For 70 characters from 21.86 tot 25.60 utterances:; for 100 characters: from 12.29 to 16.53 utterances; for 115 characters: from 9.21 to 13.07 utterances. Standard deviations and range between minima and maxima are high.

Significance

It seems that the intervention had a positive effect on the target group. Progress is more visible in the target group than in the control group, but the only significant difference is situated in the target group's long utterances above 115 characters (Wilcoxon Signed Rank Test; $p = .002$; for 100 characters as cut-off point $p = .091$; for 70 characters as cut-off point $p = .071$).

The previous table showed the results of the non-curricular discussion. We will now have a closer look at the figures for the problem solving discussion.

Table 53

Long utterances in the pre- and post-test of the problem solving discussion

Status			**Pre-test 70 characters**	**Post-test 70 characters**	**Pre-test 100 characters**	**Post-test 100 characters**	**Pre-test 115 characters**	**Post-test 115 characters**
Control group	N	Valid	15	15	15	15	15	15
		Missing	0	0	0	0	0	0
	Mean		9.53	6.6	4.27	2.6	2.93	1.73
	Std. Deviation		6.323	4.469	3.305	2.501	2.549	1.944
	Minimum		4	2	0	0	0	0
	Maximum		22	16	14	7	11	5
Target group	**N**	Valid	15	15	15	15	15	15
		Missing	0	0	0	0	0	0
	Mean		9.27	14	4.07	6.8	2.67	4.67
	Std. Deviation		5.244	3.78	3.081	3.005	2.059	2.38
	Minimum		2	9	0	1	0	1
	Maximum		18	20	10	12	6	9

Table 53 shows that, for the control group, the number of long utterances again decreases in the post-test. And again it makes no difference whether we define ‘long utterances’ as more than 70, 100 or 115 characters. For 70 characters: from 9.53 tot 6.60 utterances; for 100 characters: from 4.27 to 2.60 utterances; for 115 characters: from 2.93 to 1.73 utterances. Decrease seems the strongest when we take 100 characters as a cut-off point.

In the target group we see increasing numbers for the three cut-off points. For 70 characters: from 9.27 to 14 utterances; for 100 characters: from 4.07 tot 6.80 utterances; for 115 characters: from 2.67 to 4.67 utterances. Similar to the non-curricular discussion it appears that the intervention had a positive effect on the target group. This is especially visible when we take 115 characters as a cut-off point. Contrary to the non-curricular discussion, we observe that the average number of long utterances in the control group is more or less similar to those of the target group during the pre-test.

Significance

When we compare the scores of the pre-test with those of the post-test, we find significant progress for the target group, whichever cut-off point we chose (Wilcoxon Signed Rank Test; 100 characters: p = .016; 70 characters: p = .011; 115 characters: p = .026). Noticeable is that, more than in the non-curricular discussion, some triads produce not a single long utterance when we set the cut-off point at 100 and 115 characters. Again this can be explained by the nature of the assignment and the presence of ‘materials’ which makes conversation more implicit and turns shorter. But as mentioned earlier (see section 3.5.1 in this chapter), rapid-turntaking can be an indicator of exploratory talk (Haglund & Jeppsson, 2012; Mercer, 2000). Finally, the decrease of long utterances within the control group can be explained by the fact that more than in the target group triads went over the problem solving test rather quickly, the result of which were shorter conversations. Therefore, it is necessary we look at the proportional number of long utterances.

Proportional number of long utterances

Absolute measurements do not take into account the differences in length between all conversations. In our data, these vary between 1402 and 8143 characters. On top of that, one triad may produce 'more characters' during the pre-test than in the post-test, but the reverse is equally possible. This means that, if we want to know whether the number of long utterances evolves proportionally, we have to express these in percentages. This is shown in tables 54 and 55.

Table 54

Long utterances in the pre- and post-test of the non-curricular discussion (proportional, in percentages)

Status			**Pre-test 70 characters**	**Post-test 70 characters**	**Pre-test 100 characters**	**Post-test 100 characters**	**Pre-test 115 characters**	**Post-test 115 characters**
Control group	N	Valid	15	14	15	14	15	14
		Missing	0	1	0	1	0	1
	Mean		33.24	36.56	18.51	19.72	12.96	13,27
	Std. Deviation		11.979	7.756	10.518	7.856	8.286	5,757
	Minimum		18	24	1	10	0	5
	Maximum		52	47	34	32	28	26
Target group	N	Valid	14	15	14	15	14	15
		Missing	1	0	1	0	1	0
	Mean		38.61	37.67	21.32	24.38	15.93	19,55
	Std. Deviation		8.247	7.258	8.16	7.327	7.042	6,599
	Minimum		21	18	10	9	5	5
	Maximum		52	49	36	39	31	31

As table 54 shows, the percentage of long utterances, whether starting from 70, 100 or 115 characters, increases moderately for the control group. For 70 characters: from 33.24 to 36.56 utterances; for 100 characters: from 18.51 to 19.72 utterances; for 115 characters: from 12.96 to 13.27 utterances. Again standard deviations and the range between minima and maxima are high.

For the target group, increase is clear when we look at utterances above 100 and 115 characters. For 100 characters: from 21.32 tot 24.38 utterances; for 115 characters: from 15.93 to 19.55 utterances. There is a decrease, however, when 70 characters is taken as cut-off point, i.e. from 38.61 to 37.67 utterances. Standard deviations and the range between minima and maxima remain high.

Significance

No signifcant progress was found when comparing pre- and post-test results of the target group. The increase of long utterances above 115 characters comes closest to significance (Wilcoxon Signed Rank Test; p = .074).

Table 55 shows the percentages of long utterances for the problem solving discussion.

Table 55
Long utterances in the pre- and post-test of the problem solving discussion (proportional, in percentages)

Status			**Pre-test 70 characters**	**Post-test 70 characters**	**Pre-test 100 characters**	**Post-test 100 characters**	**Pre-test 115 characters**	**Post-test 115 characters**
Control group	N	Valid	15	15	15	15	15	15
		Missing	0	0	0	0	0	0
	Mean		17.22	17.02	8.09	6	5.52	3.624
	Std. Deviation		6.979	7.698	4.806	5.171	3.76	3.731
	Minimum		7	8	0	0	0	0
	Maximum		27	33	17	17	13	10.43
Target group	N	Valid	15	15	15	15	15	15
		Missing	0	0	0	0	0	0
	Mean		18.27	26.02	7.53	12.94	4.71	9.089
	Std. Deviation		7.485	6.821	4.939	6.206	3.414	5.199
	Minimum		8	17	0	2	0	1.47
	Maximum		33	39	17	22	10	19

Looking at the control group in table 55 we observe a decrease in long utterances for all three cut-off points. For 70 characters: from 17.22 to 17.02 utterances; for 100 characters: from 8.09 to 6 utterances; for 115 characters: from 5.52 to 3.62 utterances. In the target group we see quite the contrary. For 70 characters: from 18.27 tot 26.02 utterances; for 100 characters: from 7.53 to 12.94 utterances; for 115 characters: from 4.71 tot 9.08 utterances. Percentages of long utterances in the problem solving discussion are consistently lower than in the non-curricular discussion, which also shows by the number of zero long utterances (minima). Standard deviation and range are high for both groups.

Significance

When we compare the scores of the pre-test with those of the post-test, we find significant progress for the target group, whichever cut-off point we chose (Wilcoxon Signed Rank Test; 100 characters: p = .012; 70 characters: p = .005; 115 characters: p = .023).

Summary – length of utterances

Our results show a positive effect of the intervention for the target group, especially during the problem solving discussion: utterances are significantly longer after the intervention, which suggests more (elaborate) reasoning. Whether the cut-off point for long utterances is 70, 100 or 115 characters, the results are similar. Simultaneously, there is no progress in the control group. This suggests that target pupils have learnt to elaborate more. Of course, this does not say everything about the quality of these elaborations. For that, we must look at the fourth indicator of exploratory talk: the use of arguments.

4.2.2.4 Use of arguments

In order to analyse the quantity and quality of arguments used during the group discussions, we coded the arguments used by the pupils before and after the intervention. Coding was done separately by two researchers (Cohen's kappa = .72, i.e. good reliability). The quality of these arguments was determined by mutual agreement (inter rater reliability), using the typology of Rojas-Drummond and Zapata (2004) (see section 3.5.1 in this chapter). We will first discuss the quantity of arguments and then zoom in on their quality.

Quantity of arguments

Table 56 shows the number of each argument level and the total number of arguments during the non-curricular discussion and the problem solving conversations with Raven's Progressive Matrices for the pre- as well as for the post-test.

Table 56

Number of arguments during the non-curricular and problem solving discussion

Status			Non-curricular discussion		Problem solving discussion	
			Pre-test	Post-test	Pre-test	Post-test
Control group	N	Valid	15	14	15	15
		Missing	15	16	15	15
	Mean		13.87	10.79	30.4	22.07
	Std. Deviation		8.14	7.895	10.119	13.074
	Range		27	26	30	47
	Minimum		2	4	14	6
	Maximum		29	30	44	53
Target group	N	Valid	14	15	15	14
		Missing	13	12	12	13
	Mean		10.43	19.6	22.27	39.5
	Std. Deviation		5.626	8.484	9.859	14.453
	Range		17	27	35	51
	Minimum		4	7	6	14
	Maximum		21	34	41	65

Table 56 shows the following:

a) For the non-curricular discussion:

Control group triads use an average of 13.87 arguments before the intervention. Differences between triads are extreme, as is shown by the standard deviation (s.d = 8.14) and range (27) between minimum (2) and maximum (29). After the intervention the average number of arguments drops to 10.79. Standard deviation (7.89) and range (26) between minimum (4) and maximum (30) remain practically unchanged.

Target group triads use an average of 10.43 arguments before the intervention. The differences between triads are high, though the standard deviation (s.d = 5.62) and range (17) between minimum (4) and maximum (21) are smaller than for the control group.

After the intervention the average number of arguments rises significantly to 19.6, but standard deviation (s.d = 8.48) and range (27) increase as well and so do the minimum (7) and maximum (34).

b) For the problem solving discussion:

Control group triads use an average of 30.4 arguments before the intervention. Differences between triads are again extreme, as is shown by the standard deviation (s.d = 10.12) and range (30) between minimum (14) and maximum (44). After the intervention the average number or arguments drops to 22.07, while standard deviation (13.07) and range (47) between minimum (6) and maximum (53) increase.

Target group triads use an average of 22.27 arguments before the intervention. Here, too, the differences between triads are obvious, as standard deviation (s.d = 9.86) and range (35) between minimum (6) and maximum (41) show. After the intervention the average number of arguments increases to 39.5. Standard deviation (s.d.=14.45) and the range (51) between minimum (14) and maximum (65) also increase substantially. After the intervention the average number of arguments used during the non-curricular discussion rises significantly to 39.5. Standard deviation (s.d = 14.45) and range (51) increase as well and so do minimum (14) and maximum (65).

Significance

The increase of arguments for the target group in the non-curricular discussion is significant (Wilcoxon Signed Rank Test; p = .001). Significant differences were also found for the problem solving discussion, which shows an increase of arguments for the target group (Wilcoxon Signed Rank Test, p = .019) and a decrease for the control group (Wilcoxon Signed Rank Test, p = .050). It is noticeable that the number of arguments produced by the control group decreases almost significantly (Wilcoxon Signed Rank Test, p = .053).

Quality of arguments

The quantity of arguments may indicate progress in the use of exploratory talk but equally important is the quality of these arguments. In order to measure this the arguments in the non-curricular and problem solving discussions were also coded using the framework proposed by Rojas-Drummond and Zapata (2004): A = rudimentary argument, B = implicit argument, C = semi-explicit argument and D = explicit argument. Again two researchers coded the argument independently (Cohen's kappa = 60, i.e. reasonable reliability), after which agreement was reached through comparative discussion (see also section 3.5.1 in this chapter).

Table 57

Quality of arguments (A-D) used during the non-curricular discussion

Status			Argument level A		Argument level B		Argument level C		Argument level D	
			Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Control Group	N	Valid	15	15	15	15	15	14	15	14
		Missing	15	15	15	15	15	16	15	16
	Mean		0.27	0.07	1.27	1.6	10.4	6.64	1.93	2.36
	Std. Deviation		0.458	0.258	1.163	3.376	5.422	5.108	2.764	1.985
	Range		1	1	4	9	16	15	7	6
	Minimum		0	0	0	0	2	1	0	0
	Maximum		1	1	4	9	18	16	7	6
Target group	N	Valid	15	15	15	15	14	15	14	15
		Missing	12	12	12	12	13	12	13	12
	Mean		0	0	0.67	2.87	6.71	12.6	3	4.13
	Std. Deviation		0	0	1.113	2.588	5.413	6.812	1.71	2.066
	Range		0	0	3	8	16	20	5	7
	Minimum		0	0	0	0	1	4	0	1
	Maximum		0	0	3	8	17	24	5	8

In the control group the average use of A-arguments (0.27), B-arguments (1.27) and D-arguments (1.93) is quite low, the majority being C-arguments (10.40). Standard deviation, range and minima and maxima reveal considerable differences between the triads. After the intervention there is a slight increase of D-arguments (2.36), while the average number of C-arguments drops to 6.64, and the number of A- and B-arguments remains more or less the same. Standard deviations, ranges and minima and maxima show the same variety among triads as before the intervention.

When we look at the target group we find no A- and hardly any B-arguments before the intervention. Here, too, most arguments (6.71) are C-quality arguments, while we also find a number of D-arguments (3.00). The intervention shows change in this group: again there are no A-arguments but we found a number of B-arguments (2.87). More importantly, the number of C-arguments has nearly doubled (12.60) and the number of D-arguments has also increased (4.13). Differences between triads remain high before and after the intervention, but maxima generally increase as well.

Significance

Whether producing less or more arguments of each level, no significant differences were found in the control group. In the target group the increase of C-arguments is significant (Wilcoxon Signed Rank Test, *p = .001*).

Table 58 shows the average use of arguments, based on their quality, during the problem solving discussion.

Table 58

Quality of arguments (A-D) used during the problem solving discussion

			Argument level A		**Argument level B**		**Argument level C**		**Argument level D**	
Status			**Pre-test**	**Post-test**	**Pre-test**	**Post-test**	**Pre-test**	**Post-test**	**Pre-test**	**Post-test**
Control group	N	Valid	15	15	15	15	15	15	15	15
		Missing	15	15	15	15	15	15	15	15
	Mean		9.80	6.27	19.67	14.13	0.93	1.67	0.00	0.00
	Std. Deviation		5.388	4.559	7.394	9.319	1.100	2.289	0.000	0.000
	Range		21	13	27	37	3	9	0	0
	Minimum		3	1	7	3	0	0	0	0
	Maximum		24	14	34	40	3	9	0	0
Target group	N	Valid	15	14	15	14	15	14	15	14
		Missing	12	13	12	13	12	13	12	13
	Mean		4.40	8.29	15.60	25.86	2.20	5.29	0.07	0.07
	Std. Deviation		2.324	5.757	6.544	10.596	2.704	3.429	0.258	0.267
	Range		7	19	25	39	9	10	1	1
	Minimum		0	1	3	10	0	1	0	0
	Maximum		7	20	28	49	9	11	1	1

Table 58 shows that, in the control group, the average use of A-arguments (9.80) and B-arguments (19.67) is predominant, while C-arguments (0.93) are rare and D-arguments (0) are nonexistent. Standard deviation, range and minima and maxima once again reveal differences between the triads. After the intervention the average number of arguments decreases significantly (cf. Table 56), but there is no real quality shift, a slight increase of C-arguments (1.67) notwithstanding. Again we found no D-arguments. The variety among triads also has not changed much, as standard deviations and the ranges between minima and maxima remain considerable.

When we look at the target group we see a similar dominant pattern as in the control groups: most arguments are A-arguments (4.40) and B-arguments (15.60), while C-arguments (2.20) and D-arguments (0.07) are a minority. Looking at standard deviation, range, minima and maxima we also see the same variation among triads. After the intervention the number of arguments increases (cf. table 56). Now, we also notice a moderate quality shift: A-arguments (8.29) and B-arguments (15.60) are still a strong majority, but C-arguments (5.29) now also form a non-neglectable group. D-arguments (0.07) remain rare. While differences between groups remain high, we see a considerable upsurge of maximum scores.

Significance

While producing less arguments for the A-, B- and D- levels, and a small increase of C-level arguments, none of this is significant in the control group. In the target group the increase of C-arguments is significant (Wilcoxon Signed Rank Test, $p = .001$).

Summary – use of arguments

Quantity of arguments

As could be expected, for the non-curricular discussion the experiment made little difference in the control group, but it did make a significant difference in the target group. Not only has the average number of arguments increased in that group, but so have minimum and maximum. At the same time we notice a higher standard deviation, which suggests that some triads have increased their argumental skills more than others. In short, all triads of the target group seem to have made progress but not all have done so to the same extent.

The scores for the problem solving discussion show diffusing patterns compared with the non-curricular discussion. Different is that the average number of arguments is considerably higher than for the non-curricular discussion in both tests. This not illogical, as the number of issues to discuss and solve in Raven's test is much larger than the number of discussion points in the non-curricular discussion. In the next paragraph we will go further into this, when we discuss the quality of arguments. Taking this into account, we also see similar evolutions: high standard deviations for both control and target groups at the start, wide ranges between minima and maxima. Time seems to have a negative or at least a diffusing effect on the control group's argumental skills: the average number of arguments drops, the differences between triads increase, while at the same time the maximum score goes from 44 to 53. In the target group after the intervention all scores increase, including standard deviation. Like for the non-curricular discussion some triads seem to have made more progress than others, but significant average progress is clear whatsoever.

Quality of arguments

As we hypothesised, the quality of arguments used during the non-curricular discussion did not seem to augment in the control group, while it did in the target group. Results are in line with the scores for the quantity of arguments but the same holds for the strong differences between triads. Again, we see that triads do not start nor evolve at the same rate and pace.

As far as the quality of arguments is concerned we believe the target group has made progress, increasing its number of C-arguments compared with the control group. Comparing argument quality in the non-curricular discussion with the problem solving discussion, we cannot help noticing a difference: while C- and D-arguments are more dominant in the non-curricular discussion, A- and B-arguments dominate the problem solving discussion. This can be explained by the fact that during the Raven's test the pupils had a work sheet in front of them which they often refer to non-verbally. As a result, their arguments are characterised by more signalling words and deixis, generating more implicit (A- and B-) arguments than in the non-curricular discussions, where no work sheet is used. Nevertheless, our figures show that even then, the target group triads have raised not only the quantity but also the quality of their arguments (more C-arguments) while doing Raven's, whereas the control group triads have not.

4.2.3 Test results: problem solving skills (RQ 4)

In this section we will address RQ 4 by discussing the problem solving skills of the pupils. This skills have been measured by means of Raven's Standard/Coloured Progressive Matrices as pre- and post-test. We will first look at the scores at group level as well as the correlation between these scores and the use of key words in context (4.2.3.1). Next, we will discuss the results at individual level (4.2.3.2).

4.2.3.1 Scores at group level

a) Raven's Standard Progressive Matrices

In order to measure the problem solving skills of the pupils at group level Raven's Standard Progressive Matrices were used. All triads in both control and target groups took the test, which was split up into two units of 30 puzzles of equal difficulty (see Chapter 4, section 3.4.2).

Table 59

Scores for Raven's Standard Progressive Matrices at group level

			Pre-test	Post-test
Control group	N	Valid	30	29
		Missing	0	1
	Mean		22.83	23.03
	Std. Deviation		1.859	2.353
	Minimum		18	18
	Maximum		26	28
Target group	N	Valid	27	27
		Missing	0	0
	Mean		22.30	25.15
	Std. Deviation		2.462	1.791
	Minimum		17	22
	Maximum		27	29

On average the control group triads (n = 30 before and n = 29 after the intervention) scored 22.83 on a total of 30 points before and 23.03 after the intervention, or a .88% increase. Minimum and maximum scores for the pre-test are 18 and 26 (s.d = 1.859), and 18 and 28 (s.d = 2.353) for the post-test.

The target group triads (n = 27) started off with a slightly lower score on Raven's than the control group, i.e. 22.30 points[26], and went up to an average of 25.15 after the intervention (see figure 15), or a 12.78% increase. Minimum and maximum scores were 17 and 27 (s.d = 2.462) for the pre-test, and 22 and 29 (s.d = 1.791) for the post-test, again showing progress.

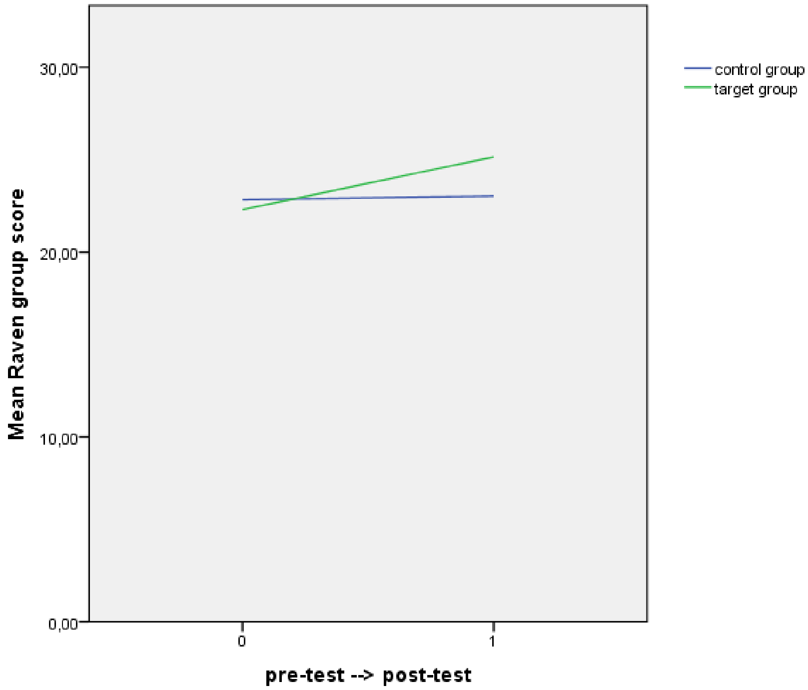


Figure 15. Average results for the problem solving test (Raven) at group level

Significance

In order to find out whether the difference between the pre-test and the post-test scores of the control group is significant, we performed a Wilcoxon Signed Rank Test which confirmed the null hypothesis. In other words, the difference between both average scores is not significant (p = .347) and at group level the control triads have not improved their problem solving skills. For the target group the same test proved the score difference between pre- and post-test to be significant (p = .000). In order to know whether the target group also did significantly better than the control group after the intervention we performed a non-parametric Mann Whitney U Test, which confirmed significancy (p = .001).

[26] Scores for Raven's Progressive Matrices increase with age (Raven, 1998). Hence, the pre-test difference of .53 points between control and target group can be explained by a difference in age: the control group pupils were – on average – six months older than the target group pupils.

b) Correlation between scores on Raven's and the use of key words in context

A number of studies explain increasing scores on Raven's by referring to evolutions in the use of key words in context viz. exploratory talk (Mercer et al., 1999; Rojas-Drummond & Zapata, 2004; Topping & Trickey, 2014; Wegerif, 1996b, etc). Through a bivariate analysis (Spearman's rho, non-parametric) we investigated this correlation and present the results in tables 60a-c.

Table 60a

Bivariate analyses: score difference for the problem solving test and the evolution of the absolute use of key words in context

			Key word difference pre-/post-tests non-curricular discussion	**Score difference pre-/post-tests problem solving**
Spearman's rho	**Score difference pre-/post-tests problem solving**	Correlation Coefficient	0.329	1
		Sig. (2-tailed)	0.087	.
		N	28	29
	Key word difference pre-/post-tests non-curricular discussion	Correlation Coefficient	1	0.329
		Sig. (2-tailed)	.	0.087
		N	29	28

We found no significant correlation between scores on Raven's and the use of key words in context comparing pre- and post-test (ρ = .329, *p = .087*), though there may be a tendency that both are related. This means that obtaining higher scores on Raven's does not significantly coincide with increased absolute use of key words. When comparing with the proportional use of key words in context, as in Table 60b, the result is more or less the same (ρ = .333, *p = .077*).

Table 60b

Bivariate analyses of the score difference for the problem solving test and the evolution of the proportional use of key words in context

			Score difference pre-/post-tests problem solving (Raven)	**Key word difference pre-/post-tests non-curricular discussion**
Spearman's rho	Score difference pre-/post-tests problem solving discussion	Correlation Coefficient	1.000	.333
		Sig. (2-tailed)	.	.077
		N	29	29
	Key word difference pre-/post-tests non-curricular discussion	Correlation Coefficient	.333	1.000
		Sig. (2-tailed)	.077	.
		N	29	30

We also looked at the correlation between the evolution of scores on Raven with the use of arguments. It appears to be rather weak (ρ = .226) and is not significant (*p = .239*).

Table 60c

Bivariate analyses of the score difference for the problem solving test and the evolution of the use of arguments

			Score difference pre-/post-tests problem solving (Raven)	**Number of arbuments. difference pre-/post-tests problem solving discussion**
Spearman's rho	Score difference pre-/post-tests problem solving (Raven)	Correlation Coefficient	1.000	.226
		Sig. (2-tailed)	.	.239
		N	29	29
	Number of arbuments. difference pre-/post-tests problem solving discussion	Correlation Coefficient	.226	1.000
		Sig. (2-tailed)	.239	.
		N	29	30

Contradictory to what we expected, the correlation between problem solving skills and the use of exploratory talk as expressed by an increased use of arguments and in-context key words, is rather weak. Therefore, a regression analysis (Tables 61a-c) had to show whether the increased use of key words has an influence on the evolution of scores on Raven's whatsoever.

Table 61a

Regression analysis of the score difference for the problem solving test and the use of the proportional use of key words in context during the problem solving discussion

Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**
1	.386[a]	.149	.117	2.070

a. Predictors: (Constant), Delta-selprct-rpm-EFF-NUL

The model shows a .386 correlation with the evolution of the scores on Raven's and significantly explains 14.9% of this evolution (*p = .039*), as the anova shows:

Table 61b

ANOVA[a]

Model		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
1	Regression	20.260	1	20.260	4.726	.039[b]
	Residual	115.740	27	4.287		
	Total	136.000	28			

a. Dependent Variable: score difference pre-/post-test RPM groups

b. Predictors: (Constant). Delta-selprct-rpm-EFF-NUL

Table 61c

Coefficients a						
		Unstandardised Coefficients		Standardised Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	1.786	.397		4.499	.000
	Proportional use of key words. difference pre-/post-tests problem solving discussion	.276	.127	.386	2.174	.039

a. Dependent Variable: scoreverschil NUL-EFF RPM groepen

The influence of the independent variable key words in context on the evolution of the scores on Raven's is positive (Beta = .386) and significant (*p = .039*).

Summary – problem solving skills at group level

At group level, experiment triads improved their problem solving skills significantly compared to the control group and compared to the beginning of the experiment. Also, the increasing use of key words in context positively influences the evolution of scores on the problem solving test, though based on earlier studies we had expected a stronger effect. We believe high standard deviations may explain the difficulty to generalise results.

4.2.3.2 Scores at individual level

In order to measure the problem solving skills of the pupils at individual level an extended version of Raven's Coloured Progressive Matrices was used (see section 3.5.2 in this chapter). Table 62 shows the pre- and post-test scores (maximum score for each test = 30).

Table 62
Scores for Raven's Coloured Progressive Matrices at individual level

Status			Pre-test	Post-test
control group	N	Valid	90	90
		Missing	5	5
	Mean		20.78	21.69
	Std. Deviation		3.197	2.986
	Minimum		5	15
	Maximum		27	29
target group	N	Valid	80	80
		Missing	3	3
	Mean		21.36	22.38
	Std. Deviation		2.645	3.054
	Minimum		14	15
	Maximum		27	29

On average the pupils of the control group (n = 90) scored 20.78 on a total of 30 points for the pre-test and 21.69 for the post-test, or a 4.38% increase. Minimum and maximum scores are 5 and 27 (s.d = 3.197) for the pre-test and 15 and 29 (s.d = 2.986) for the post-test. Comparatively, the pupils of the target group (n = 100) scored 21.36 for the pre-test and 22.38 for the post-test, or a 4.78% increase. Minimum and maximum scores are 14 and 27 (s.d = 2.645) for the pre-test and 15 and 29 (s.d = 3.054) for the post-test. Figure 16 presents the score evolution of both the control and the target group pupils.

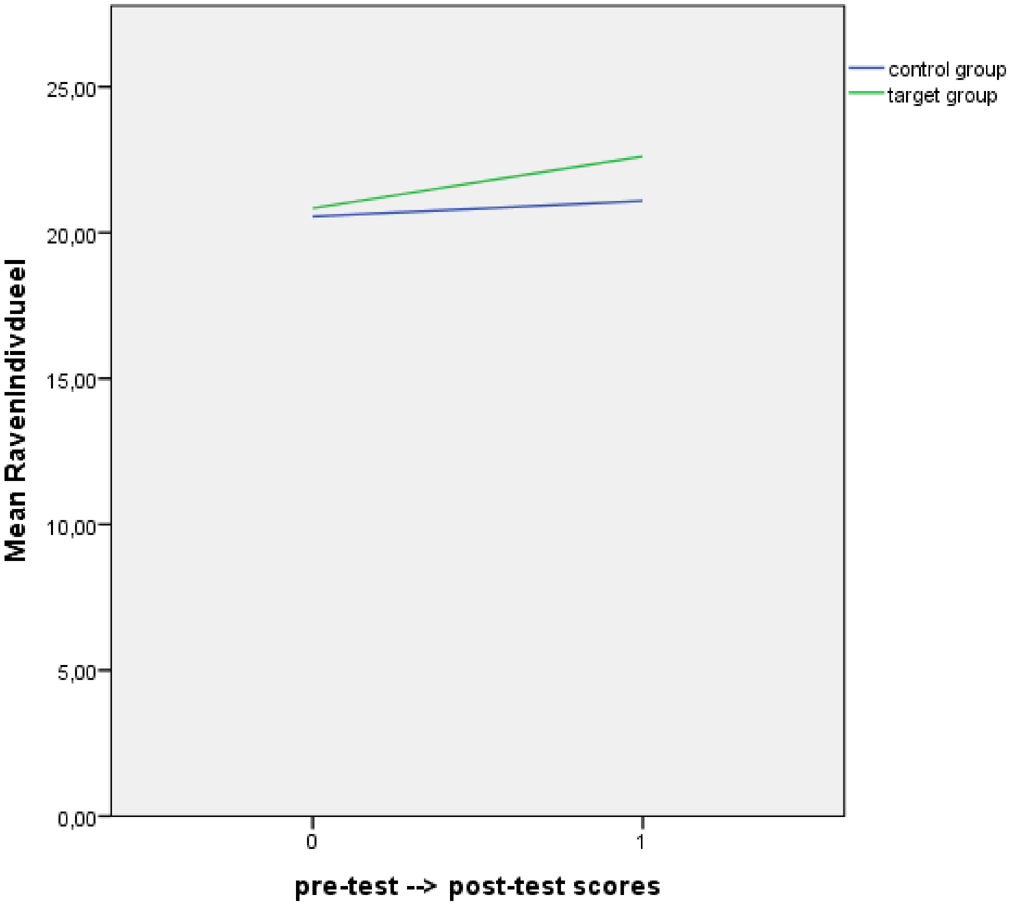


Figure 16. Average individual results for the problem solving test

We performed a parametric Paired Samples T Test to determine progress significancy on Raven's, which was confirmed for the target group (*p = .010*) but rejected for the control group (*p = .097*). We also performed an Independent Samples T Test to determine significancy for the score difference between control and target group after the intervention, but none was found.

Summary – problem solving skills at individual level

Based on our analysis we conclude that pupils of the target group have significantly improved their problem solving skills. The pupils of the control group have improved as well but not significantly. The difference in progress between control and target group is not significant either, however. Contrary to the results at group level, it is noticeable that the average starting score for pupils of the target group is higher than that of the pupils of the control group. Most importantly, however, average group scores exceed average individual scores.

4.2.4 Impact of individual variables (RQ 5)

In this section we want to discuss the influence of independent variables on the use of exploratory talk, more specifically the influence of characteristics of individual pupils on the use of key words in context. We will first discuss influences on the non-curricular discussion (4.2.4.1) and then look at influences on the problem solving discussion (4.2.4.2).

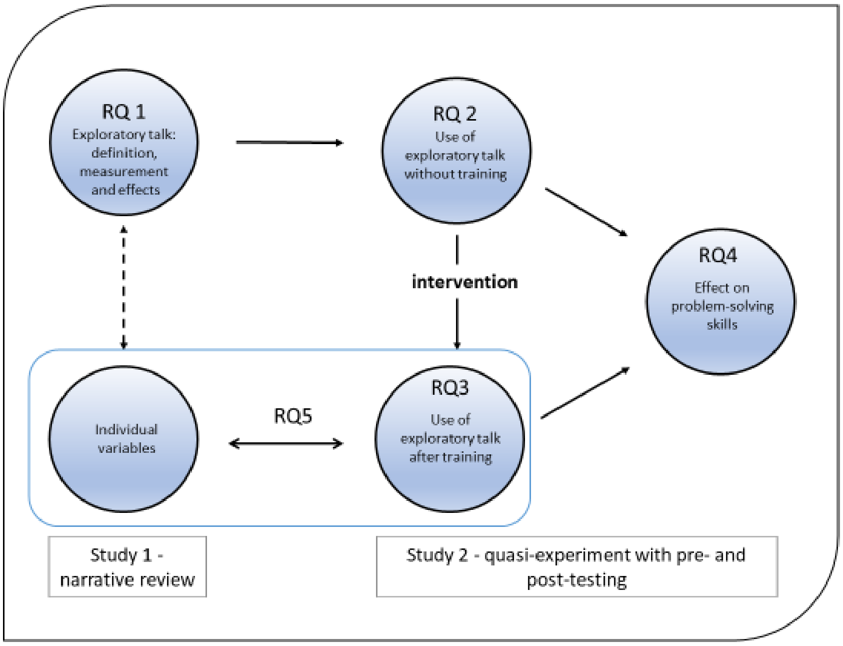


Figure 17. Research Design: focus on RQ 5 in relation to RQ 3

Study 1 and the background information for Study 2 provided us with independent variables to be measured, i.e. gender, GOK indicators[27], spelling (LVS spelling) and mathematics (LVS mathematics) skills[28]. The influence of each variable was examined separately. In order to find the correlations between key word use and individual variables we took two steps. First, we identified and counted the use of key words in context for each individual pupil within a triad. Second, we applied a linear mixed effects model[29] in order to determine the impact of the independent variables, and estimate the interaction effects between these variables and differences between pre- and post-test scores and between control group and target group.

4.2.4.1 Non-curicular discussion

At group level we saw an increase of the use of key words in context (see section 4.2.2.1) after the intervention. As could be expected, table 63 shows similar results at individual level.

Table 63

Average use of key words (absolute numbers) in context during the non-curricular discussion

Estimates of Fixed Effects[a]							
Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	41.600	7.067	30.432	5.884	.000	27.170	56.030
EFF1-NULL0	-9.578	6.894	29.575	-1.389	.175	-23.666	4.510
EG1-CG0	-1.761	9.940	31.344	-0.177	.860	-22.026	18.503
EFF1-NULL0 * EG1-CG0	26.900	9.748	30.456	2.760	.010	7.004	46.795

a. Dependent Variable: SOM-LM.
EFF1-NULL0 = main effect post-test
EG1-CG0 = main effect target group
EFF1-NULL0 * EG1-CG0 = two-way effect post-test * target group

Table 63 shows that the control group uses an average of 41.6 key words in context during the pre-test and 9.6 key words in context less during the post-test. This decrease, although

[27] GOK indicatoren or 'indicatoren gelijke onderwijskansen' (equal opportunities for education) were introduced in Flemish education in 2001 as indicators for extra individual support. Schools receive additional financial means to organise education for pupils who meet criteria like low economic family status, limited language skills, etc. In the course of this study GOK was renamed SES (socioeconomic status), sharing the same goals. In this study we consistently refer to GOK.
[28] LVS or 'leerlingvolgsysteem' comprises a number of validated tests, analytical instruments and pedagocial or didactic suggestions for technical reading skills, spelling and mathematics. They have been developed between 1997 and 2006 for primary education (Dudal, 2000, 2001).
[29] As linear mixed model statistics are more suitable to deal with missing values (we did not receive all information about mathematic skills and language skills) we preferred this approach to repeated measures ANOVA. Moreover, it allows to take into account the multilevel structure of the data (individual pupils are nested in groups) and the fact that the variance between pupils and between groups does not have to be identical on both occasions.

somewhat surprising, is not significant (*p = .175*). The target group starts off with 39.9 key words in context, which is almost as much as the control group (*p = .860).* But after the intervention the use of key words in context by the target group increases signicantly with 26.9 units (*p = .010*).

As not all conversations have the same length, we repeated these calculations to determine the proportional use of key words in context. In other words, we divided the total number of key words in context by the number of words per conversation, the results of which are shown in table 64.

Table 64
Average proportional use of key words in context during the the non-curricular discussion

Estimates of Fixed Effects[a]							
Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	0.045	0.006	29.877	7.069	.000	0.032	0.058
EFF1-NULL0	-0.010	0.006	28.689	-1.585	.124	-0.024	0.003
EG1-CG0	-0.018	0.009	30.308	-1.95	.061	-0.036	0.001
EFF1-NULL0 * EG1-CG0	0.025	0.009	29.146	2.83	.008	0.007	0.043

a. Dependent Variable: PROP-LM-TOTAAL.
EFF1-NULL0 = main effect post-test
EG1-CG0 = main effect target group
EFF1-NULL0 * EG1-CG0 = two-way effect post-test * target group

Table 64 is more difficult to interpret as the 'Estimate' column is the result of a calculation which does not render an absolute number but a proportion of key words in context. (Multiplied by 100 these proportional numbers can be seen as percentages.) Nevertheless we see similar results as in table 63. The control group's use of key words in context still decreases non-significantly (*p = .124*). The difference in key word in context use by the control and target group during the pre-test is larger than when we consider only absolute numbers of key words in context and even close to being significant (*p = .061*): figures show that the target group uses proportionally far less key words in context during the pre-test. After the intervention the proportional use of key words in context by the target group increases significantly (*p = .008*). At the same time, the difference in 'Estimate' scores between control group and target group is smaller than when we only looked at absolute numbers (cf. control group: 0.045 - 0.010 = 0.035 'Estimate'; target group: 0.045 - 0.018 + 0.025 = 0.052 'Estimate'; difference between both groups: 0.052 - 0.035 = 0.017 'Estimate').

Knowing this, we used similar statistical procedures to test correlations with four independent variables: gender, GOK, LVS spelling and LVS mathematics. Table 65 shows the results of significance calculations of four statistically relevant items considering the absolute number of in-context key words used.

Whenever we speak of key words in this section and if not specified, we refer to key words in context.

Table 65

Results of linear mixed effect models of 'in-context key word (absolute number) use' with gender, GOK, spelling level, mathematics level during the non-curricular discussion as interaction effects

Absolute number of key words					
Independent variable		**Gender**	**GOK**	**Spelling**	**Maths**
		(1 = boys)	(1 = GOK)	(1 = high)	(1= high)
	Intercept*	49.484	53.627	40.628	43,737
Main effects	**Independent variable**	-14.191	-12.886	2.186	-4.182
		(p = .092)	(p = .475)	(p = .802)	(p = .633)
	Target group	-1.339	-16.529	-7.250	-15.883
		(p = .910)	(p = .495)	(p = .535)	(p = .197)
	Post-test	-14.803	-12.469	-12.618	-8,900
		(p = .091)	(p = .512)	(p = .125)	(p = .297)
Two-way effects	**Independent variable * target group**	-1.949	15.910	8.949	21.638
		(p = .866)	(p = .508)	(p = .475)	(p = .081)
	Independent variable * post-test	9.405	3.097	6.841	-1.325
		(p = .308)	(p = .870)	(p = .459)	(p = .889)
	Target group * post test	35.013	31.460	34.775	25,538
		(p = .005)	(p = .214)	(p = .004)	(p = .045)
Three-way effects	**Independent variable * post-test * target group**	-14.704	-4.942	-15.374	3.370
		(p = .254)	(p = .845)	(p = .247)	(p = .801)

** intercept in model = control group pre-test, zero value for independent variables*

Table 65 must be interpreted as follows:

Intercept:

The intercept can be read as the expected number of key words used if all independent variables have a zero value. In this table it is represented by the figures of key words used by the control group during the pre-test. For the model in which we include the effect of gender this means expected key word use in the control-group at the pre-test of girls is 49.48. For the effect of GOK, spelling skills and mathematics skills this means: non-GOK pupils (53.63 key words), pupils with low spelling skills (40.63 key words) and pupils with low mathematics skills (47.74 key words). Main effects will be compared with these figures.

Main effects:

- **Independent variable**, compared with the intercept: this is the number of key words used by respectively boys, GOK pupils, pupils with high spelling skills and pupils with high mathematics skills during the pre-test. Key words use by these pupils is lower, except for pupils with high spelling skills (+2,19 key words), but none of the differences are significant. A tendency towards significance shows for boys, who use 14,19 key words less than girls ($p = .092$). All in all, this means that within the control group there are no significant differences between these categories of pupils at the pre-test.

- **Target group**, compared with the intercept: figures indicate the differences in the amount of in-context key words used by respectively girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills in the target group during the pre-test. Compared with the control group these figures are negative, which indicates a lower key word use than in the control group, but differences are not significant. So, at the start of the intervention no significant differences were found between control and target group.

- **Post-test**, compared with the intercept: during the post-test and in the control group less key words are used by girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills than during the pre-test, but no significant differences were found. A tendency towards significance shows for girls, who use 14.8 key words less than during the pre-test ($p = .091$). We can conclude that pupils in the control group did not change their amount of key word use between pre- and post-test significantly.

Two-way effects:

- **Interaction independent variable – target group:** these estimates indicate whether the difference between control and target group at the pre-test is dependent on the variables introduced earlier (gender, GOK, spelling level, mathematics level). For instance, the difference in key word use between target and control group at the pre-test is slightly larger (-1.949) for boys than for girls, but this is not statistically significant. Again, there are no significant differences, though a tendency towards significance shows for pupils with high mathematics skills (+21,64 key words, $p = .081$). This basically means that the differences between control and target group at the pre-test are not dependent on gender, GOK or spelling and mathematics levels.

- **Interaction independent variable – post-test:** these estimates quantify whether the differences between pre- and post-test in the control group are dependent on the independent variables. For instance, where we saw that girls in the control group use less key words between pre- and post-test (-14.8, see main effect post-test) for boys this is 9.4 less. Nevertheless, these interactions are all statistically insignificant. This means that the differences between pre- and post-test in the control group are not dependent on gender, GOK or spelling/mathematics skills.

- Interaction target group – post-test: these estimates show whether the intervention influences the difference between pre- and post-test for the target group. If we look at the model with gender, we learn that the difference between pre- and post-test is significantly different for girls in the target group (35.0 more key words) than for the same category of pupils in the control group. In the target group girls increase their key word use with 20.2 units (= -14.8 + 35.0). Similar differences show when we consider the other independent variables in the model, but as far as key word use by non-GOK pupils is concerned, the difference with the pre-test control group is not significant.

Three-way effect:

- **Interaction independent variable – post-test - target group:** finally, a 3-way interaction is modelled, indicating whether the effect of the intervention on evolution between pre-and post-test can be explained by the analysed independent variables. This is not the case for any of the variables, for none of the differences are significant. This leads us to the conclusion that, when we consider in-context key word use in absolute numbers, the intervention has the same effect on boys and girls, GOK- and non-GOK pupils, and pupils with high as well as low spelling/mathematics skills. All target pupils make progress in the use of in-context key words after the intervention, but this progress is not significant considering the independent variables.

In table 66 similar figures have been processed, but this time we only consider the proportional use of in-context key words.

Table 66

Results of linear mixed effect models of 'in-context key word (proportional number) use' with gender, GOK, spelling level, mathematics level during the non-curricular discussion as interaction effects

Proportional number of key words					
Independent variable		**Gender (1 = boys)**	**GOK (1 = GOK)**	**Spelling (1 = high)**	**Maths (1= high)**
	Intercept*	.054	.063	.043	.045
Main effects	**Independent variable**	-.017	-.010	.004	-.0002
		(p = .022)	(p = .579)	(p = .619)	(p = .975)
	Target group	-.021	-.045	-.019	-.025
		(p = .054)	(p = .046)	(p = .064)	(p = .024)
	Post-test	-.150	-.390	-.012	-.009
		(p = .052)	(p = .020)	(p = .098)	(p = .220)
Two-way effects	**Independent variable * target group**	.005	.018	.004	.012
		(p = .593)	(p = .430)	(p = .727)	(p = .250)
	Independent variable * post-test	.009	.011	.005	-.001
		(p = .226)	(p = .513)	(p = .508)	(p = .897)
	Target group * post test	.032	.051	.030	.023
		(p = .004)	(p = .022)	(p = .005)	(p = .041)
Three-way effect	**Independent variable * post-test * target group**	-.014	-.015	-.012	.003
		(p = .200)	(p = .525)	(p = .302)	(p = .782)

** intercept in model = control group pre-test, zero value for independent variables*

Table 66 shows more significant effects than whene we considered absolute numbers of key words use only. We will zoom in on these significancies:

Intercept:

As in table 65 the intercept includes the figures of in-context key words used by the control group during the pre test: girls (.054 or 5.4%), non-GOK pupils (.063 or 6.3%), pupils with low spelling skills (.043 or 4.3%) and pupils with low mathematics skills (.045 or 4.5%). Main effects will be compared with these figures.

Main effects:

- **Independent variable**, compared with the intercept: this is the number of in-context key words used by respectively boys, GOK pupils, pupils with high spelling skills and pupils with high mathematics skills during the pre-test. The tendency towards significance shown for boys in table 65 (*p = .092*), has been turned into full significance *(p = .022).* So, key word use is significantly lower for boys, but non-significantly lower for non-GOK pupils and nearly equal for pupils with high mathematics skills. It is non-significantly higher for pupils with high spelling skills (+.004 or 0.4%, *p = .619*). This means that within the control group there is one significant difference between pupils at the pre-test, i.e. when gender is taken into account.

- **Target group**, compared with the intercept: figures indicate the differences in the amount of in-context key words used by respectively girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills in the target group during the pre-test. Compared to the control group, these figures are negative, which indicates lower key word use than in the control group. The differences are significant considering GOK *(p = .046*) and mathematics skills *(p = .024)* and nearly significant for gender *(p = .052)* and spelling skills *(p = .064*). So, at the start of the intervention significant differences show between control and target group considering GOK and mathematics as independent variables. Lower use is strong among girls and pupils with low spelling skills, but it is not significantly lower.

- **Post-test**, compared with the intercept: during the post-test and in the control group less key words are used by girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills. The difference is strong (.390 or -39%) and significant *(p = .020*) for GOK, strong (0.150 or 15%) and nearly significant (*p = .052*) for gender (girls) and less outspoken (0.012 or 1,2%) and with a tendency towards significance for spelling level. We can conclude that non-GOK pupils in the control group used significantly less key words after the intervention, while gender has a strong effect.

Two-way effects:

- **Interaction independent variable – target group:** these estimates indicate whether the differences between control and target group at the pre-test depend on the independent variables. Boys, GOK pupils, and pupils with high spelling/mathematics skills use more key words than girls, non-GOK pupils and pupils with low spelling/mathematics skills, but as there are no significant differences, we can conclude that differences between control and target group at the pre-test are not dependent on gender, GOK or spelling and mathematics levels.

- **Interaction independent variable – post-test:** these estimates show whether the differences between pre- and post-test in the control group can be explained by the independent variables. Girls, non-GOK pupils, pupils with low mathematics skills use less key words while pupils with low spelling skills use more, but none of these interactions are statistically significant. So, the differences between pre- and post-test in the control group are not dependent on gender, GOK or spelling and mathematics levels.

- Interaction target group – post-test: these estimates show whether the intervention influences the difference between pre- and post-test. All differences are significant. This means that when we consider the independent variables in the model, boys, GOK pupils, and pupils with high spelling/mathematics skills in the target group use more key words than the same categories of pupils in the control group.

Three-way effect:

- **Interaction independent variable – post-test - target group:** this interaction model indicates whether the effect of the intervention on evolution between pre-and post-test can be explained by the analysed independent variables. This is not the case for any of the variables. So, considering context key word use, the intervention has the same effect on boys and girls, GOK- and non-GOK pupils, and pupils with high as well as low spelling/mathematics skills. Progress in key words use (absolute numbers) as shown in table 65 is confirmed when we look at their proportional use.

4.2.4.2 Problem solving discussion

At group level we saw an increase of the use of key words in context (see section 4.2.2.1). As can be expected, table 67 shows similar results at individual level.

Table 67

Average use of key words (absolute numbers) in context during the problem solving discussion

Estimates of Fixed Effects[a]							
Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	56.956	6.471	29.882	8.801	0.000	43.738	70.174
EFF1-NULL0	-24.022	5.957	29.097	-4.033	0.000	-36.204	-11.841
EG1-CG0	-10.134	9.103	30.823	-1.113	0.274	-28.705	8.436
EFF1-NULL0 * EG1-CG0	34.894	8.42	30.066	4.144	0.000	17.7	52.087

a. Dependent Variable: SOM-LM.
EFF1-NULL0 = main effect post-test
EG1-CG0 = main effect target group
EFF1-NULL0 * EG1-CG0 = two-way effect post-test * target group

Table 67 shows that the control group uses an average of 57 key words in context during the pre-test and 24 key words in context less during the post-test. This decrease is significant (*p = .000*). The target group starts off with (56.96 – 10.13 =) 46.8 key words in context, which is less than the control group, but the difference between both groups is not significant (*p = .274*).

After the intervention the use of key words by the target group increases signicantly with 34.9 units (*p = .000*).

Again, we divided the total number of key words in context by the number of words per conversation to obtain the proportional use of key words in context. Table 68 shows the results:

Table 68
Average proportional use of key words in context during the problem solving discussion

Estimates of Fixed Effects[a]							
Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	0.060	0.005	31.768	11.179	0.000	0.049	0.071
EFF1-NULL0	-0.029	0.006	31.001	-5.163	0.000	-0.034	-0.017
EG1-CG0	-0.028	0.008	32.379	-3.655	0.001	-0.043	-0.012
EFF1-NULL0 * EG1-CG0	0.037	0.008	31.559	4.761	0.000	0.021	0.054

a. Dependent Variable: PROP-LM-TOT.
EFF1-NULL0 = main effect post-test
EG1-CG0 = main effect target group
EFF1-NULL0 * EG1-CG0 = two-way effect post-test * target group

As in table 67 the 'Estimate' column in table 68 is the result of a calculation which does not render a number of key words in context but a proportional figure. Most noticeable is that all differences are now significant: the control group's use of key words in context decreases after the intervention (*p = .000*); the target group uses significantly less key words in context during the pre-test than the control group (*p* = .001) and after the intervention the proportional use of key words in context by the target group increases significantly.

As the evolution of key word use seems stronger in the problem solving conversation we considered the chance for independent variables to be influential realistic, especially for mathematics, as it requires similar skills as the solving of Raven's. To that end, table 69 was constructed in the same way as table 66 and shows the following figures:

Table 69

Results of linear mixed effect models of 'in-context key word (absolute number) use' with gender, GOK, spelling level, mathematics level during the problem solving discussion as interaction effects

Absolute number of key words Independent variable		**Gender** (1 = boys)	**Gok** (1 = GOK)	**Spelling** (1 = high)	**Maths** (1= high)
	Intercept	54.018	54.858	57.420	47.856
Main effects	**Independent variable**	5.287	2.248	-1.040	17.803
		(p = .513)	(p = .894)	(p = .899)	(p = .033)
	Target group	-1.060	-16.230	-16.040	-7.069
		(p = .923)	(p = .474)	(p = .135)	(p = .520)
	Post-test	-22.438	-17.019	-25.358	-16.210
		(p = .004)	(p = .308)	(p = .001)	(p = .026)
Two-way effects	**Independent variable * target group**	-17.186	6.763	8.619	-10.964
		(p = .128)	(p = .764)	(p = .466)	(p = .344)
	Independent variable * post-test	-2.851	-7.504	3.006	-15.285
		(p = .726)	(p = .652)	(p = .714)	(p = .065)
	Target group * post test	30.164	20.587	35.355	24.113
		(p = .006)	(p = .353)	(p = .001)	(p = .024)
Three-way effects	**Independent variable * post-test * target group**	8.884	15.562	1.079	21.841
		(p = .436)	(p = .483)	(p = .927)	(p = .063)

** intercept in model = control group pre-test, zero value for independent variables*

Table 69 must be interpreted as follows:

Intercept:

The intercept is made up by the figures of in-context key words used by the control group during the pre-test. For the effect of the independent variables gender, GOK, spelling skills and mathematics skills this means: girls (54.02), non-GOK pupils (54.86), pupils with low spelling skills (57.42) and pupils with low mathematics skills (47.86). Main effects will be compared with these figures.

Main effects:

- **Independent variable**, compared with the intercept: this is the number of key words used by respectively boys, GOK pupils, pupils with high spelling skills and pupils with high mathematics skills during the pre-test. Key word use by these pupils is slightly lower for pupils with high spelling skills (-1.04 key words), higher for boys and GOK-pupils and significantly higher for pupils with high mathematics skills (+17.8 key words, *p = .033*). Essentially, this means that within the control group there are no significant differences between the four categories of pupils at the pre-test, except for the category of mathematics skills.

- **Target group**, compared with the intercept: figures indicate the differences in the amount of key words in context used by respectively girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills in the target group during the pre-test. Compared with the control group, these figures are all negative and reveal a lower to much lower key word use before the intervention in the target group. None of these differences are significant, though. We conclude that, at the start of the intervention, there are no significant differences between the categories of pupils in the control vs. the target group.

- **Post-test**, compared with the intercept: during the post-test and in the control group less in-context key words are used by girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills. These differences are significant for all categories of pupils, but not for non-GOK pupils. We conclude that pupils in the control group significantly changed their amount of key words use between pre- and post-test in a negative way and that this change can be explained by differences in gender and spelling/mathematics skills.

Two-way effects:

- **Interaction independent variable – target group:** these estimates indicate whether the difference between control and target group at the pre-test can be explained by the independent variables. The figures show a mixed pattern: boys and pupils with high mathematics skills use less key words after the intervention, GOK-pupils and pupils with high spelling skills use more. In both cases, however, no significant differences were found. This means that the differences between control and target group at the pre-test do not depend on gender, GOK or spelling/mathematics skills.

- **Interaction independent variable – post-test:** these estimates quantify whether the differences between pre- and post-test in the control group can be explained by the independent variables. All interactions statistically non-significant, except for a tendency towards significance concerning mathematics skills, which implies that pupils with high mathematic skills use considerably less key words after the intervention. This means that the differences between pre- and post-test in the control group are not dependent on gender, GOK or spelling, but can perhaps be explained by the mathematics level of the pupils.

- Interaction target group – post-test: these estimates show whether the intervention influences the difference between pre- and post-test. If we look at the model with all independent variables, we learn that all target pupil groups use more key words than pupils in the control group. The difference between pre- and post-test is not significant for non-GOK pupils (+20.59, *p = .353*) and significant for girls (30.16 words, *p = .006*), pupils with low spelling skills (+35.36, *p = .001*) and pupils with low mathematics skills (+24.11, *p = .024*).

Three-way effect:

- **Interaction independent variable – post-test - target group:** this interaction shows whether the effect of the intervention on evolution between pre-and post-test can be explained by the analysed independent variables. This is not the case for any of the variables, though a tendency towards significance shows for pupils with high mathematics skills, who increase their key word use with 21.84 units (*p = .063*). So, when we consider context key words use, the intervention has the same effect for gender, GOK and LVS spelling and virtually the same effect on pupils with high mathematics skills. On the whole, all target pupils use more of in-context key words after the intervention, but this evolution is not significant when effects of independent variables are taken into consideration.

In table 70 similar figures have been obtained but this time we only consider the proportional use of in-context key words.

Table 70

Results of linear mixed effect models of 'in-context key word (proportional number) use' with gender, GOK, spelling level, mathematics level during the problem solving discussion as interaction effects

Proportional number of key words					
Independent variable		**Gender**	**Gok**	**Spelling**	**Maths**
		(1 = boys)	(1 = GOK)	(1 = high)	(1= high)
	Intercept*	.062	.059	.059	0,047
Main effects	**Independent variable**	-.003	.001	.001	.025
		(p = .712)	(p = .931)	(p = .845)	(p = .001)
	Ttarget group	-.025	-.028	-.030	-.018
		(p = .010)	(p = .179)	(p = .002)	(p = .052)
	Post-test	-.031	-.027	-.028	-0,020
		(p = .000)	(p = .074)	(p = .000)	(p = .003)
Two-way effects	**Independent variable * target group**	-.005	.0003	.005	-.020
		(p = .660)	(p = .990)	(p = .665)	(p = .055)
	Independent variable * post-test	.004	.001	-.001	-.017
		(p = .558)	(p = .937)	(p = .863)	(p = .024)
	Target group * post-test	.037	.032	.035	.027
		(p = .000)	(p = .114)	(p = .000)	(p = .007)
Three-way effects	**Independent variable * post-test * target group**	.001	-.006	.004	.020
		(p = .945)	(p = .774)	(p = .696)	(p = .053)

** intercept in model = control group pre-test, zero value for independent variables*

Table 70 must be interpreted as follows:

Intercept:

The intercept is represented by the percentages of in-context key words used by the control group during the pre-test. For the effect of the independent variables gender, GOK, spelling skills and mathematics skills this means: girls (.062 or 6.2%), non-GOK pupils (.059 or 5.9%), pupils with low spelling skills (.059 or 5.9%) and pupils with low mathematics skills (.047 or 4.7%). Main effects will be compared with these figures.

Main effects:

- **Independent variable**, compared with the intercept: this is the percentage of key words used by respectively boys, GOK pupils, pupils with high spelling skills and pupils with high mathematics skills during the pre-test. Key words use by these pupils is nearly equal for all categories of pupils, the differences being non-significant. Exception must be made for pupils with high mathematics skills, whose percentage of key words is significantly higher than that of pupils with low mathematics skills (.025 or 2,5%, *p = .001*). This means that within the control group there are no significant differences between the four categories of pupils at the pre-test, execept for the category of mathematics skills.

- **Target group**, compared with the intercept: figures indicate the differences in the percentages of key words used by respectively girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills in the target group during the pre-test. Compared with the control group, these figures are all negative, which indicates a lower in-context key word use in the target group than in the control group, before the intervention. Moreover, these differences are significant for girls and pupils with low spelling skills, and nearly significant for pupils with low mathematics skills. So, at the start of the intervention significant negative differences show in three out of four categories of pupils of the target group compared to the control group.

- **Post-test**, compared with the intercept: during the post-test and in the control group less key words are used by girls, non-GOK pupils, pupils with low spelling skills and pupils with low mathematics skills. These differences are significant for girls, pupils with low spelling skills and pupils with low mathematics skills, and nearly significant for non-GOK pupils. We conclude that pupils in the control group significantly changed their amount of key word use between pre- and post-test in a negative way and that this change can be explained by the independent variables gender, spelling and mathematics skills.

Two-way effects:

- **Interaction independent variable – target group:** these estimates indicate whether the difference between control and target group at the pre-test can be explained by the independent variables. Boys and pupils with high mathematics skills use less key words, GOK-pupils and pupils with high spelling skills use more key words. In both cases, no significant differences were found, except near significant difference for pupils with high mathematics skills (*p = .055).* This means that the differences between control and target group at the pre-test do not depend on gender, GOK or spelling, but may be explained by mathematics levels.

- **Interaction independent variable – post-test:** these estimates show whether the differences between pre- and post-test in the control group can be explained by the independent variables. All interactions are all statistically non-significant, except for mathematics skills, where we see that pupils with high skills use significantly less key words (-.017 or -1.7%, *p = .007*). This means that the differences between pre- and post-test in the control group are not dependent on gender, GOK or spelling, but can be explained by mathematics skills.

- Interaction target group – post-test: these estimates show whether the intervention influences the difference between pre- and post-test. If we look at the model with all independent variables, we learn that all target pupil groups use more key words than pupils in the control group. The difference between pre- and post-test is significant for girls and for pupils with low spelling/mathematics skills. Non-GOK pupils have also increased their key word use after the intervention and compared with the control group, but not significantly. So, the individual increase of key words use by the target group can be explained by three out of four independent variables, i.e. gender, spelling and mathematics.

Three-way effect:

- **Interaction independent variable – post-test - target group:** this interaction shows whether the effect of the intervention on evolution between pre- and post-test can be explained by the analysed independent variables. This is not the case for any of the variables, though for mathematics skills differences are nearly significant (.020 or 2%, $p = .053$). So, when we consider context key words use, the intervention has the same effect on boys and girls, GOK- and non-GOK pupils, pupils with high and low spelling skills but perhaps a different, positive effect on pupils with high mathematics skills. On the whole, all target pupils, except GOK pupils, use more of in-context key words after the intervention, but this evolution is not significant.

As mathematics skills reveal several significant differences in our models, we believe this is worth further examination.

All control group pupils use significantly less key words in context after the intervention, the largest decline showing for pupils with high mathematics skills. Target group pupils, on the other hand, use significantly more key words in context after the intervention. Both pupils with low and high mathematics skills show progress. This is visualised in figure 18.

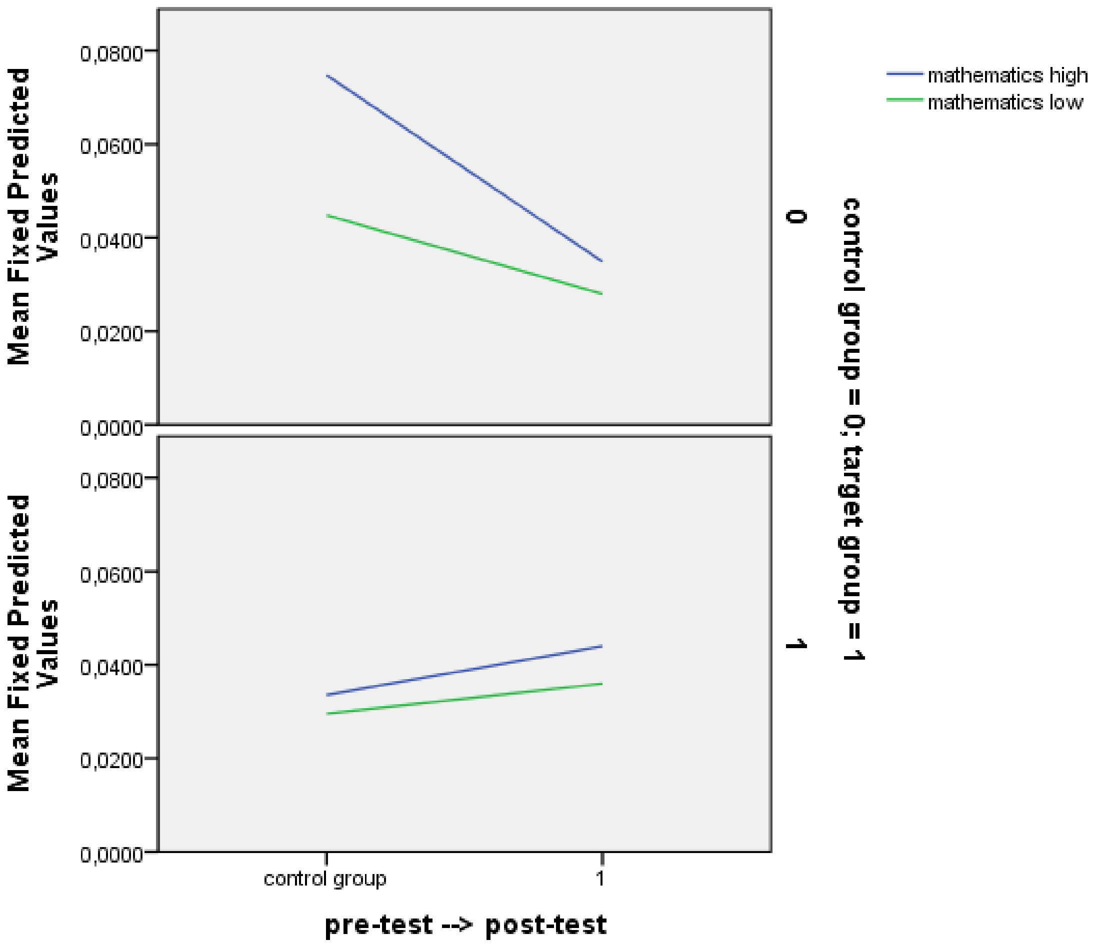


Figure 18. Influence of mathematics level on in-context key words use in the problem solving discussion

We used the same statistics to find out whether the impact of the level of mathematics is visible in the use of all categories of key words in context or only in some. Therefore we replaced the dependent variables absolute and proportional number of key words by their respective categories of key words, which relate to the six ground rules (see Chapter 4, section 3.5.1). For instance, category 1 comprises words that indicate the organisation of ideas within each argument with conjunctions like 'because (of)', 'and', 'also', 'if' … ' which are generally used to connect a claim or a challenge with one or more arguments and/or a conclusion. Category 3 comprises questions, which connects this category to the first. The results are shown in table 71 and 72:

Table 71

Correlation of categories of key words use (absolute number) with mathematics skills during the problem solving discussion

Absolute number of key words		**Key words category 1**	**Key words category 2**	**Key words category 3**	**Key words category 4**	**Key words category 5**	**Key words category 6**
Control group*	**Pre-test**	11.34 (*p* = .003)	1.39 (*p* = .607)	.50 (*p* = .000)	-.0896 (p = .597	.85 (*p* = .221)	4.38 (*p* = .069)
	Post-test	-10.93 (*p* = .007)	.13 (*p* = .967)	-.56 (*p* = .003)	.0086 p = .972)	-1.14 (*p* = .161)	-3.25 (*p* = .271)
Target group	**Pre-test**	-9.530 (*p* = .074)	-.54 (*p* = .886)	-.3892 (*p* = .057)	.1218 (*p* = .617)	.27 (*p* = .788)	-1.55 (*p* = .644)
	Post-test	15.94 (*p* = .006)	2.65 (*p* = .556)	.5873 (*p* = .031)	.1901 (*p* = .579)	.86 (*p* = .468)	2.97 (*p* = .482)

Table 72

Correlation of categories of key words use (proportional number) with mathematics level during the problem solving discussion

Proportional number of key words		**Key words category 1**	**Key words category 2**	**Key words category 3**	**Key words category 4**	**Key words category 5**	**Key words category 6**
Control group*	**Pre-test**	.0139 (*p* = .000)	.0036 (*p* = .162)	.0004 (*p* = .001)	-.00008 (*p* = .953)	.0011 (*p* = .069)	.0057 (*p* = .010)
	Post-test	-.0117 (*p* = .002)	-.0009 (*p* = .742)	.0005 (*p* = .002)	-.00001 (*p* = .993)	-.0009 (*p* = .190)	-.0036 (*p* = .132)
Target group	**Pre-test**	-.0124 (*p* = .016)	-.0033 (*p* = .370)	.0003 (*p* = .074)	-3.3370 (*p* = .987)	-.0005 (*p* = .591)	-.0034 (*p* = .262)
	Post-test	.0151 (*p* = .005)	.0024 (*p* = .557)	.0005 (*p* = .031)	.00007 (*p* = .785)	.0008 (*p* = .410)	.0024 (*p* = .474)

Both tables 71 and 72 show that the influence of mathematics level is significant or nearly significant for the use of key words in context of category 1 and 3 in both control and target groups and in both pre- and post-tests. So, the effects of mathematics skills on the use of key words in context and hence, exploratory talk, seems to be restricted to two categories of key words resp. ground rules. One involves questions like ‘What do you think’ and ‘Why do you think that?’, the other involves pupils’ answers to those questions, i.e. justification by argument building. Apparently, in problem solving conversations, pupils have learnt to not only probe for arguments and challenge them but also to formulate arguments in response.

4.2.5 Overall summary of the main study

Looking at the main study, we see that the intervention had many positive and significant effects in the target group which do not occur in the control group. This section summarise s our findings, relating them to our research questions.

Indicators of exploratory talk (RQ 2-3)

Tables 73 and 74 show to what extent both the control and target group have (significanlty) evolved from pre- to post-tests when considering the four indicators of exploratory talk. The abbreviations and symbols in the tables should be read as following:

ncd	non-curricular discussion
rpm	problem solving discussion (Raven’s Progressive Matrices)
=	no change after the intervention
+	increase after the intervention
-	decrease after the intervention
A-D	quality of arguments (A = rudimentary, B = implicit, C = semi-explicit, D = explicit)
p	(near) significance
(n.s.)	no significance

Table 73

Progress of control and target group after the intervention: key words in context and turn-taking

	Key words in context				**Turn-taking**			
	% key words in context per conversation		**% key words in context in proportion to the whole conversation**		**Quantitative symmetry**		**Interactive dominance/recession**	
	ncd	**rpm**	**ncd**	**rpm**	**ncd**	**rpm**	**ncd**	**rpm**
Control group	=	- (n.s.)	- (n.s.)	- (n.s.)	=	=	=	=
Target group	=	+ (*p* = .004)	+ (n.s.)	+ (*p* = .001)	+ (*p* = .059)	+ (n.s.)	=	+ (*p* = .034)

Table 74

Progress of control and target group after the intervention: long utterances and use of arguments

	Long utterances (100 characters)		Arguments			
			Quantity		Quality	
	ncd	**rpm**	**ncd**	**rpm**	**ncd**	**rpm**
Control group	- (n.s.)	- (n.s.)	- (p = .053)	- (p = .050)	A-; B+, C-, D+ n.s.	A-; B-; C+; D= n.s.
Target group	+ (p = .091)	+ (p = .016)	+ (p = .001)	+ (p = .001)	A=; B+; D+ (n.s.) C+ (p = .001)	A+; B+; D+ (n.s.) C+ (p = .001)

Tables 73 and 74 show that most significant, positive effects are visible in the problem solving conversation. The control group did not evolve significantly for any of the four indicators of exploratory talk. For key words in context there was a general, though not significant decrease, while for turn-taking there was no evolution at all. The number of arguments also decreased, on the verge of significancy for the non-curricular discussion and significantly for the problem solving discussion. The only progress was seen in the quality of certain levels of arguments, though here, too, no scores were significant.

Contrastively, the target group shows significant increase of indicators in 8 of 14 cases and a positive tendency in 2 cases. Most progress was noted during the problem solving discussion (6 of 8 significant cases). No decrease of indicators was seen.

Problem solving skills (RQ 4)

Table 75 summarises the group and individual scores on the problem solving pre- and post-tests for the control and target group.

Table 75

Scores for the problem solving tests

Status	**Triads**		**Individual pupils**	
	Pre-test	**Post-test**	**Pre-test**	**Post-test**
Control group	22.83	23.03 (n.s.)	20.78	21.69 (n.s.)
Target group	22.30	25.15 (p = .000)	21.36	22.38 (p = .010)

Statistically speaking, the pupils of the target group significantly increased their scores for the problem solving test, both at group and individual level, whereas no significant progress was noted in the control group.

Impact of individual variables (RQ 5)

The following tables show the correlations between key words use and the individual independent variables in the non-curricular discussion (table 76) and in the problem solving discussion (table 17). In table 76 correlations prove to have a (near) significant influence.

Table 76

Significant correlations between key words use and gender resp. mathematics in the non-curricular discussion

Independent variable			**Gender (1 = boys)**	**Gok (1 = GOK)**	**Spelling (1 = high)**	**Maths (1= high)**
	Intercept					
Main effects	**Independent variable**	absolute	-14.19 (p= .092)			
		proportional	-.017 (*p* = .022)			
	Target group	absolute				
		proportional	-.021 (*p* = .054)	-.045 (*p* = .046)	-.019 (*p* = .064)	-.025 (*p* = .024)
	Post-test	absolute	-14.80 (*p* = .091)			
		proportional	-.150 (*p* = .052)	-.390 (*p* = .020)	-.012 (*p* = .098)	
Two-way effects	**Independent variable * target group**	absolute				21.64 (*p* = .081)
		proportional				
	Independent variable * post-test	absolute				
		proportional				
	Target group * post-test	absolute	35.01 (*p* = .005)		34.78 (*p* = .004)	25.54 (*p* = .045)
		proportional	.032 (*p* = .004)	.051 (*p* = .022)	.030 (*p* = .005)	.023 (*p* = .041)
Three-way effects	**Independent variable * post-test * target group**	absolute				
		proportional				

If we look at the significant figures (*p* = .050 of less) we conclude that:

- for gender: boys of the control group use proportionally less key words than girls at the start of the experiment, but after the intervention the difference is no longer significant. Girls of the target group use more key words after the intervention than girls of the control group. As we discussed earlier, boys of the target group also make progress but less than girls, but the difference in progress between boys and girls is not significant.

- for GOK: non-GOK pupils of the target group use less key words at the start of the experiment than non-GOK pupils of the control group. Also, non-GOK pupils of the control group use less key words after the intervention, whereas those of the target group make significant progress. As tables 65 and 66 showed, GOK-pupils of the target group use more key words than the control group but there is no significant difference in progress with non-GOK pupils.

- for spelling skills: pupils with low spelling skills of the target group use more key words after the intervention than the control group. Pupils with high spelling skills also make progress but this is less outspoken than pupils with low spelling skills. Yet, no significance between both categories of pupils was found.

- for mathematics skills: pupils with low mathematics skills of the target group use more key words after the intervention than the control group. Pupils with high mathematics skills make nearly similar progress. As for spelling skills, no significance between both categories of pupils was found.

Table 77 shows the correlations between the use of key words and the independent variables in the problem solving discussion.

Table 77

Significant correlations between key words use and gender resp. mathematics in the problem solving discussion

Independent variable			Gender (1 = boys)	Gok (1 = GOK)	Spelling (1 = high)	Maths (1= high)
	Intercept					
Main effects	**Independent variable**	absolute				17.80 (*p* = .033)
		proportional				.025 (*p* = .001)
	Target group	absolute				
		proportional	-.025 (*p* = .010)		-.030 (*p* = .002)	-.018 (*p* = .052)
	Post-test	absolute	-22.44 (*p* = .004)		-25.36 (*p* = 001)	-16.21 (*p* = .026)
		proportional	-.031 (*p* = .000)	-.027 (*p* = .074)	-.028 (*p* = .000)	-.020 (*p* = .003)
Two-way effects	**Independent variable * target group**	absolute				
		proportional				-.020 (*p* = .055)
	Independent variable * post-test	absolute				-15.29 (*p* = .065)
		proportional				-.017 (*p* = .024)
	Target group * post-test	absolute	30.16 (*p* = .006)		35.36 (*p* = .001)	24.11 (*p* = .024)
		proportional	.037 (*p* = .000)		.035 (*p* = .000)	-.027 (*p* = .007)
Three-way effects	**Independent variable * post-test * target group**	absolute				21.84 (*p* = .063)
		proportional				.020 (*p* = .053)

If we look at the significant figures (*p* = .05 or less) we conclude that:

- for gender: girls of the target group use proportionally less key words than girls of the control group at the start of the experiment. Further, girls of the control group use less key words after the intervention than at the pre-test. Girls of the target group show significant progress in key words use after the intervention. Boys of the target group make slightly more progress (cf. table 70) but the difference is non-significant.

- for GOK: non-GOK pupils of the control group use less key words after the intervention than at the start of the experiment. No significant differences with GOK-pupils were found.

- for spelling skills: target group pupils with low spelling skills use less key words than the same category of control group pupils at the start of the experiment. Control group pupils with low spelling skills use less key words after than before the intervention. Target group pupils with low as well as high spelling skills significantly enhance their key word use. Progress of both categories of pupils is not significant.

- for mathematics skills: control group pupils with high mathematics skills use more key words than pupils with low mathematics skills. Also, though only nearly significant, target group pupils with low mathematics skills use less key words at the start of the experiment than control group pupils. After the intervention, control group pupils with low ánd high mathematics skills use less key words than before the intervention. After the intervention, all target pupils increase their use of key words. The progress target pupils make is significant and in favour of pupils with high mathematics skills. For target pupils, progress in key word use, as was shown in tables 71 and 72, is significant for category 1 and category 3 key words, which focus on asking for and formulating arguments.

Chapter 5: Summary and conclusions

In this chapter we will summarise our findings, draw conclusions and answer the research questions. The first section zooms in on the context of the dissertation (1), after which methodology (2) and results (3) are discussed.

1. The context of the study

Over the last twenty years a line of mostly British researchers has repeatedly demonstrated that during collaborative activities pupils can generate more learning outcomes if they are intensively trained in the use of exploratory talk. The notion of exploratory talk was launched by Barnes (1976) and defined, worked out and empirically grounded by Mercer (Mercer, 1995; Wegerif & Mercer, 1997a). Pupils can master exploratory talk if they learn and respect the following ground rules:

1. All relevant information is shared;
2. The group seeks to reach agreement;
3. The group takes responsibility for decisions;
4. Reasons are expected;
5. Challenges are acceptable;
6. Alternatives are discussed before a decision is taken;
7. All in the group are encouraged to speak by other group members.

If pupils neglect those rules or only sparsely use them, their conversations are likely to be dominantly cumulative or disputational. These types of conversation may have certain value in e.g. brainstorms (cumulative) or debates (disputational), but their effect on collaborative learning is fairly limited, if not negative. In a classroom context or in any context in which people are to learn from one another, exploratory talk is preferable; it has the strongest educational potential. Unfortunately, none of us is born with these exploratory skills, we have to learn them at school and we have to do so as early as in primary school. As Hart and Risley (1995) found, performance in secondary education depends on the mastery of these skills in primary education.

In Flemish and Dutch education the concept of exploratory talk for learning has long been (and still is) unknown among teachers and educators. Flemish curricula, school books and publications, as well as teacher educators, do promote collaborative activities in the classroom, but as far as group work is concerned, most focus on organisation and structure, on the allocation of duties and on ideas for assignments. Conversational nature, needs and requirements are either neglected or not dealt with systematically, let alone based on a scientifically sound frame of reference. The ground rules for exploratory talk are either not mentioned at all or have to be deduced from various sections in the sources we have mentioned. We have no knowledge of a pedagogy that stimulates the consequent and effective learning and use of exploratory talk in a transparent and systematic way. If we add to this the observation that research on speaking and listening skills in mother tongue education has been scarce in Flanders as well as in the Netherlands for decades, the purpose of this study must be clear. The results of international research on exploratory talk in the classroom have been so convincing that we were anxious to know whether exploratory talk would also be teachable and learnable, and generate similar learning effects in Flemish classrooms.

2. Methodology

In order to measure learning effects of exploratory talk in pupil-pupil conversations, Wegerif and Mercer (1997a), Mercer et al. (1999), Rojas-Drummond et al. (2003), Rojas-Drummond and Zapata (2004), Rojas-Drummond et al. (2006) and other researchers set up various empirical experiments in which target groups were first taught the basic principles (ground rules) of exploratory talk via basic lessons, while control groups did regular group work without the exploratory dimension. After the basic lessons followed a usually extended period of regular and systematic collaborative activities (group work) during which feedback and feed forward reflection stimulated pupils to master exploratory talk and apply it to different school subjects, mostly science and mathematics. Pre- and post-testing was organised to determine learning effects, often including measurements of the development of problem solving skills.

From the beginning it was our plan to replicate this research design, but scientific literature soon pointed out that much had happened since Mercer started his first experiments in schools in the nineties. Relevant questions were: is Mercer's (1995) definition still standing or has the concept been redefined, deepened, elaborated, etc? How was, respectively is exploratory talk being measured and what were the results? Which independent variables are currently being taken into account? And has research methodology changed over the last twenty years? In order to answer these questions we decided to study literature systematically and integrate this study into a narrative review (Study 1). For that purpose we consulted six international scientific databases and Google Scholar. We searched the databases for the term 'exploratory talk' in combination with 'learning', 'pupil', 'teacher' and other relevant educational terms. This way, 115 articles were selected for close reading and coding based on a set of questions and subquestions. Coding was done using NVivo 10.

Based on this literature study, we decided to stay close to Mercer's (1995) original research design, as this would be the first of its kind to be undertaken in Flemish education and because his research methodology, the sociocultural discourse analysis, was repeatedly applied in other replicator studies. That is, we decided to conduct a quasi-experiment with pre- and post-testing in primary schools, followed by qualitative and quantitative discourse analysis (Study 2). For data analysis we also added parameters of the Rojas-Drummond et al. (2004) study.

The quasi-experiment involved five Antwerp primary schools, totalling 11 classes, 11 teachers and 198 (or 163 unique)[30] pupils. Because of the methodological, organisational and technical complexity of the experiment we started with a pilot class, the profile of which differed from the other classes, as these were 4th form pupils and most of them had a low socioeconomic background and weaker language skills. The pilot study (one class) preceded the main study (10 classes) for which five control and five target groups of 5th and 6th form pupils were defined.

The pilot study lead to a number of adjustments of the main study: the basic lessons and certain lesson materials were adapted to become more appealing and challenging; a checklist was made which would help the teacher to construct group work which stimulated exploratory talk in a consistent way; the use of the prompts for reasoning became less scripted (pupils were also allowed to use them after the basic lessons whenever they felt the need); evaluation sheets on the mastery of ground rules were introduced at two levels, i.e. self and peer evaluation, in order to boost feedback and feed forward reflections; the application of the problem solving test (Raven's Progressive Matrices) was altered in order to avoid a ceiling effect; the propositions of the non-curricular discussion were 'sharpened' and we included GOK as an independent variable to determine individual impact on the use of exploratory talk.

In each class the experiment took 14 weeks, including two weeks for pre- and post-testing. In the target groups the five basic lessons took four weeks and were followed by eight weeks of group work. In the control groups the experiment took 14 weeks as well: two weeks for pre- and post-testing and 12 weeks for regular group work.

During the experiment pupils of all classes were divided into groups of three or maximum four pupils by the teachers. In each class three triads were closely examined, totalling 15 in the control group and 15 in the target group: of every group work at least ten minutes were devoted to conversation, with minimal interruption by the teacher. These conversations were video and audio recorded, and transcribed. Apart from that all pupils (57 triads/groups) took part in all classroom activities, including pre- and post-testing. As in Mercer's (1995) study, control groups were not taught exploratory talk but immediately started doing group work, at least twice a week. Target groups first learnt the ground rules of exploratory talk via five basic lessons given by the teachers who had previously been introduced to the matter and trained for classroom practice. These lessons were adaptations of the ones described in Mercer and Littleton (2007). Every lesson consisted of practical assignments for which pupils had to work together in groups (triads), of whole-class moments, and of feedback and feed forward reflections. After these basic lessons the children did group work, twice a week, including reflective activities about their progress in handling the ground rules.

[30] Through the application of switching replication 35 pupils took part both as control and as target group. See Chapter 4, sections 3.2 and 4.2.1.3.

The experiment in the pilot school taught us that it was important for every group work assignment to pursue a similar set of educational goals which focused on problem solving. For that reason the teachers of the main study were all given a checklist which helped them to include the problem solving angle in their group work lessons. We also asked the teachers to mail their lesson plan to us beforehand so we could provide any necessary feedback and make suggestions for adjustments.

Pre- and post-testing was organised in two ways: one was a group discussion about a non-curricular topic. The other was a problem solving test (Raven's Progressive Matrices, Standard and Adapted Coloured version) which was taken both in groups and individually. Each test was divided into two parts of equal difficulty (30 puzzles).

Conversations were transcribed verbatim and stored for qualitative as well as quantitative analysis. For this study we analysed all pre- and post-test conversations, i.e. the non-curricular discussion and the problem solving test at group level. Analysis comprised four indicators of exploratory talk: the use of key words in context, turn-taking, long utterances and the number and quality of arguments. We hypothesised that pupils would significantly use more key words in context in the post-test. Also, turn-taking would be more democratic, utterances would be longer as pupils would elaborate more on claims they made and the quantity and quality of arguments would augment. This would mean exploratory talk is 'learnable'. As an additional learning effect both the group and individual scores on Raven's would be significantly higher after the intervention, which means they would have increased their problem solving skills. Finally, we investigated the influence of certain independent variables on the use of key words in context at an individual level. Literature and previous research suggest possible influences of gender, socioeconomic status (referred to as GOK in this study) and mathematics skills. To this we added language skills, though due to a lack of background information this had to be restricted to only one aspect, i.e. spelling skills.

For the analysis of our data we employed the sociocultural approach, which combines quantitative and qualitative methods. We used NVivo 10 to count all key words for exploratory talk that appeared in the conversations. These are words and word groups like 'I think', 'agreed', 'Why?', Because…', 'Would you…', etc, which were categorised based on their primary use in the ground rules, for which we used the framework presented by Rojas-Drummond and Zapata (2004). Because a rude count of these key words is not sufficient (pupils use them outside an exploratory context as well), qualitative analysis was done to determine whether these key words were used in the proper – exploratory – context, hence 'key words in context'. Key words that were not used in context were marked and left out of statistical analysis. Qualitative analysis also revealed the use of key words we had not anticipated on, e.g. 'Wait' to express a thinking process. These key words were, in turn, quantitatively analysed on their occurrence.

In order to collect information about turn-taking and long utterances, all relevant conversations were processed in a spreadsheet programme which was also used to determine and count the use of key words in context at an individual level. Information about the number and quality of arguments was obtained by analysing each conversation qualitatively. This was done separately by two researchers who discussed the quality of arguments in order to reach agreement. All results were processed with a spreadsheet programme.

Finally, statistical analyses were performed to generate frequency tables, find significant correlations and causal relationships, and to determine the influence of the independent variables on the use of key words in context.

3. Results

We will now address the results of our study and answer the research questions.

RQ 1. What is exploratory talk, how is it measured and which effects does it have?

In order to answer these research questions, we conducted a literature study in the form of a narrative review of 115 peer reviewed articles selected from six electronic scientific databases and Google Scholar. Search focused on the concept of 'exploratory talk' in an educational context and was limited to the years 1976-2016. In 1976, after studying pupil-pupil conversations during group work, Barnes (1976, 2010) launched the notion of exploratory talk as a means for learning. Building on Barnes' work Mercer (1996) formulated a definition which contains the following characteristics: talk is critical and constructive; statements and suggestions are made for joint consideration; challenges and counterchallenges are justified; alternative hypotheses are offered; knowledge is made publicly accountable and reasoning is more visible in the talk; joint agreement is strived for. These qualities distinguish exploratory talk from disputational and cumulative talk, which have less educational value.

Definition

Mercer's definition has been quoted or referred to in nearly all articles reviewed. Though it has been subject to certain elaborations and refinings, it remains valid and is commonly being used by other researchers. We found (near) synonyms of exploratory talk, such as *exploratory discourse, collaborative argumentation/reasoning, dialogic talk, transactive reasoning*, and also *accountable talk* and *critical discussion*. Some of these synonyms go back to earlier education or communication theory, others are meant as improvements or as attempts to label this type of talk more accurately. Over the last fifteen years some researchers have deepened and widened the concept, introducing elaborating or refining terms like *incipient* and *elaborate exploratory* talk (Rojas-Drummond et al., 2003), *inclusive* and *exclusive exploratory talk* (Rajala et al., 2012), and *reflective* and *operational talk* (Brevig, 2006). Again, depending on theoretical choices, overarching terms have been proposed, such as *IDRF* (Wegerif & Mercer, 1997a), *co-constructive talk* (Rojas-Drummond et al., 2006), *critical learning* (Riley, 2006), *collaborative* *argumentation* (Golanics & Nussbaum, 2008) and *dialogic reason* (Dourneen, 2013). We concluded that the notion of exploratory talk has become part of a widely acknowledged triad taxonomy of classroom talk: disputational – cumulative – exploratory. At the same time, the process of precisely labelling what has been called exploratory talk has been going on for decades (including its variants) and still awaits termination.

Based on the studies of Rajala et al. (2012) on turn-taking and interactive dominance/recession and of Polo et al. (2015) on group identity as important conditions for and the result of exploratory talk, we decided to elaborate on Mercer's definition and describe it as

> *a specific form of co-constructive interaction expressed through discourse*, in which partners *equally participate to maintain a sustained focus on a shared line of reasoning*. They engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counterchallenged, but challenges are justified, and alternative hypotheses are offered. Knowledge is made publicly accountable. *Reasoning is more elaborate* and more visible in the talk. Progress then emerges from the eventual joint agreement reached *and supported by group ownership*.

Measurement

In order to describe how exploratory talk is measured we did a close reading of 88 studies involving empirical research and comparative analyses, methodology and theoretical issues. We found no methodological consensus as far as measuring exploratory talk is concerned. Literature revealed almost as many qualitative as mixed methods studies, while only two studies were purely quantitative. About one third of the 88 studies used Mercer's sociocultural discourse analysis, which combines qualitative and quantitative analytical techniques. Nearly two thirds of the empirical experiments took place in primary education, focusing on pupil-pupil talk. Interventions often included the teaching of ground rules of exploratory talk and pre- and post-testing to measure effects. Most studies chose science and mathematics as their experimental school subjects. Indicators of exploratory talk which were analysed most often are the use of key words (in context), turn-taking, long utterances (elaboration) and the quality and quantity of arguments.

Learning effects

In general, effects have been found in four domains: language, psychology, cognition and learning. Many studies formulated effects in more than one domain, which suggests that effects are interactive or interwoven, and that they can trigger or amplify one another. For clarity's sake we will discuss them separately.

Language effects are: increased use of key words in context, the use of more and better arguments, longer utterances, an increased use of open-ended questions and better expression of opinions.

Psychological effects include that interactions become more symmetrical (pupils give and take speaking turns in a democratic way), the focus on self-identity shifts to group identity or group ownership and to the need for consensus. Also, silent pupils take up more assertive roles, pupils are more open to ideas developed by others or to ideas found in other sources, the feeling of self-efficacy increases, collaborative attitudes (equity, active listening, goal-mindedness) grow and cultural divides are more easily crossed.

Cognitive effects comprise improved academic performance in general. Specifically, problem solving skills and critical thinking increase. Also noted is the stronger development of thought and meaning. Accordingly, many studies support Vygotsky's claim that social learning precedes and stimulates individual learning.

Learning gains - which brings us to the pedagogical domain - have been found during collaborative activities for various school subjects: science and mathematics, geography, reading comprehension and writing skills in language classes, and music education. Reasoning skills improved during philosophy projects and computer-based assignments, suggesting the transferral qualities of exploratory talk. Last but not least, pupils also improved their metacognitive skills.

RQ 2. To what extent do pupils of the third level (primary school) use exploratory talk in group assignments?

Our literature study has made clear that exploratory talk is quite complex and cannot be described based on the relative occurrence of one indicator alone. It is more than the mere use of key words in context. It also has to do with democratic turn-taking, longer utterances denoting elaboration and the length and quality of arguments. These four criteria are mentioned in most empirical studies, which is the main reason we decided to use them as well. Other indicators, e.g. group identity, turned out to be hard to quantify.

We will now discuss each indicator in more detail.

Key words in context

Due to a lack of benchmarks it is not easy to answer this question. We did not encounter any study which puts out marks on a continuum before or behind which a conversation can be labelled exploratory, cumulative or disputational, based on one criterion or the other. Nevertheless, it is plausible to assume that conversations which contain a lot of key words in context will be more exploratory than conversations which contain few or none (Mercer, 1995; Wegerif & Mercer, 1997a; Mercer et al., 1999, etc).

At the beginning of our experiment in the pilot school it became clear that pupils did not respect the ground rules of exploratory talk much. This was reflected very well in their language: neither in the non-curricular nor in the problem solving discussions did we find evidence of all markers or exploratory talk. Markers like 'What do you think?', 'Why do you say…?', 'I agree' or the short 'agreed', 'I think that …' were absent, while markers like 'because…', indicating the beginning of an argument for a claim, were rare.

In the main study, featuring a majority of children with good language skills, we found more markers in the pre-test conversations than in the pilot study, but in most triads the average use of key words in context remained low. Especially markers like 'What do you think?', 'Why do you …?', 'I think/believe that…' and 'I agree with' or 'agreed' were only used now and then. Some triads used many markers but did not use them in (exploratory) context.

Turn-taking

In the pilot study turn-taking was reasonably symmetrical in the non-curricular discussions. We observed this in the pre- as well as in the post-test. Also, during both tests there was no interactive dominance or recession. The problem solving discussion told a different story: all triads showed quantitative asymmetry and interactive dominance and recession occurred in one of the three triads.

In the main study quantitative asymmetry was found in multiple conversations, as were interactive dominance and recession. Some conversations showed more symmetry, as one or more pupils expressed the need to 'talk orderly', but after close analysis it appeared that the initiative to talk exploratively and implicitly use the ground rules only came from one pupil in the group. Especially in the problem solving discussions, democratic turn-taking was scarce at the start of the experiment. Additionally, quantitative asymmetry dominated both control and target groups during both pre-test discussions.

Long utterances

Long utterances were rare in the pilot group. In the non-curricular discussion we found 13 utterances longer than 100 characters in triad 1, 8 in triad 2 and only 1 in triad 3. In the problem solving discussion long utterances were virtually non-existent. We found 1 in triads 1 and 2, and 3 in triad 3. The lack of benchmarks makes it somewhat tricky to assess these figures, but it is obvious that the number of long utterances was low. As many pupils of this group had weak language skills, this is not surprising.

In the main study the length of the conversations differed substantially. Therefore, we calculated the proportional use of long utterances and analysed this as well. In the non-curricular discussion 18,5% of all utterances of the control triads and 21,3% in the target group triads exceeded the cut-off point of 100 characters. The problem solving discussion showed lower figures: 8% for the control group and 7,5% for the target group. Again, it is difficult to assess these figures, as there are no benchmarks. We did notice that both target and control groups started off somewhat equally, as we expected they would. The different percentages for the non-curricular and the problem solving discussion can be explained by the nature and organisation of the discussion: during the non-curricular discussion pupils did not have any lesson materials in front of them, except for a small leaflet that contained three propositions to be discussed. During the problem solving discussion pupils had a set of 30 different puzzles in front of them which they had to solve and which caused much more non-verbal (pointing) and implicit conversation.

Arguments

In the pilot study the average number of arguments was fairly low. In the non-curricular discussion, which lasted ten minutes, triad 1 formulated 6 arguments, triad 2 formulated 15 arguments and triad 3 produced 9. More arguments appeared in the problem solving discussion: 11 (triads 1 and 2) and 4 (triad 3). Video recordings showed a lot of non-verbal behaviour which inhibited or replaced verbal argument building. More importantly, we hardly found any high quality arguments, neither in the non-curricular nor in the problem solving discussion. Most arguments were rudimentary and implicit, which reflects the weaker language skills of this group. Also, arguments were scattered across conversations instead of forming a coherent line of argumentation. In other words, if an argument was formulated, it was rarely followed by a counterargument or by further elaboration.

In the main study, control triads formulated 2 to 29 (and an average of nearly 14) arguments during the non-curricular discussion. Target triads formulated 4 to 21 (and an average of 10,4) arguments. The difference may be somewhat surprising, but then again, discussion about moral topics (like: ‘Rich people should donate half of their money to the poor’) is a rather divergent assignment. It can lead discussions many ways, including cumulative and disputational talk, in which ‘the interactants may often end up maintaining their own opinion, even after listening to those of the other participants’ (Tin, 2003, p. 54). This implies that the number of arguments may still be large, but constructive reasoning and consensus building based on argumentation are not guaranteed.

During the problem solving discussion control triads produced 14 to 44 (and an average of some 30) arguments. Target triads produced 6 tot 41 (and an average of 22) arguments. Here, differences between both groups may be more surprising, as Raven’s Progressive Matrices can be considered as a convergent assignment. In such assignments only one outcome is expected or is true, while the pupils need to pursue a single goal (Tin, 2003). Also, some triads went through the puzzles very fast, while hardly exploring possibilities, whereas others took their time and examined each puzzle in a more critical way. This, too, may explain the enormous differences between groups.

RQ 3. To what extent do pupils use exploratory talk after a 12 week training?

In order to answer this question we analysed the post-test conversations in the same way as we did for the pre-test conversations, and compared both results. Again, we looked at the use of key words in context, turn-taking, long utterances and the quantity and quality of arguments.

We did not have a control group to match the pilot group, but nevertheless it is interesting to see what the intervention did to this group.

Looking at the main study, we conclude that the intervention had many positive effects in the target group. These effects are mostly significant for the problem solving conversation. Pupils have not only mastered the ground rules, they have even shown proof of meta-thinking by e.g. reprimanding one another when one pupil violated one of the ground rules.

We will now discuss each indicator in more detail and include points of view found in our literature study.

Key words in context

After the intervention/training, the three triads of the pilot group used significantly more key words in context during both types of discussions. We also noticed that the pupils used some markers abundantly after the intervention, while not having used them in the pre-test at all. This is especially the case for the why-question and the consensus seeking 'agreed'. It is feasible to conclude that the post-test conversations of these pupils had become much more exploratory than the pre-test ones. The pupils had learnt to abide by the rules and doing so, they had acquired some of the key words almost as if they were second language vocabulary. As a negative side effect, though, they did not always use these words in the proper context.

In both the pilot and main study we skipped all key words that were out of context via qualitative analysis. We made further calculations on that in the main study. There, we noticed that in some conversations a mere 5% of the key words had to be skipped after the intervention, as they were out of the exploratory context, whereas in others this was nearly 20%. This suggests that some triads had made a lot of progress, while others had not. It may also suggest, however, that some triads already used a considerable amount of key words in context from the beginning, which left them less room for progress.

On average, especially in the problem solving discussions, the triads of the target groups in the main study used significantly more key words in the required exploratory context than before the intervention, while the control groups made little or no progress. The difference is also significant when comparing the number of key words in context of the target group with those of the control group. This is comparable to the findings of Wegerif, Mercer, and Dawes (1999) who saw all key features of exploratory talk increase significantly after their intervention in the target group.

When we consider all the words of a conversation, pupils in the target group used roughly 8 to 18% key words in the non-curricular conversation and 7 to 15% in the problem solving conversation. This is substantially more than what Wegerif (1996a) found on average, but then we included more key words in our data, while the difference between languages (one language may use more words than the other to express the same content) and cultural differences regarding (classroom) conversation may also play a role as far as the amount of key words is concerned.

Progress in the use of key words in context was less outspoken in the non-curricular discussion than in the problem solving discussion. We see parallels with the Rojas-Drummond et al. (2006) study which found that 'engaging in explicit and accountable reasoning […] was useful for success in joint solving of the reasoning test where the aim was to find the single correct underlying essentially mathematical pattern that united a series of pictures' (Rojas-Drummond et al., 2006, p. 92). The researchers further found that explicit reasoning did not serve more divergent tasks as good as convergent tasks, an explanation which is also offered by Tin (2003). As a discussion has more divergent features and consensus is not always required, pupils may very well - either consciously but most probably unconsciously - lose the feeling of necessity to use exploratory talk.

Our study shows very interesting parallels with the study made by Wegerif, Mercer and Dawes (1999) who found that, after the intervention, the use of key words in context by the control groups had decreased while solving Raven's. We encountered the same phenomenon. As a possible explanation for this Wegerif, Mercer and Dawes (1999) hypothesised that the triads 'found the task and situation less engaging the second time around than the first' (Wegerif, Mercer & Dawes, 1999, p. 511) which may have reflected on their language use. As we noticed that control triads went through their post-test faster than many target triads (we measured the time they spent doing the tests), we tend to agree with this. But another explanation which we take into consideration is suggested by Webb et al. (2016). They found different results for Raven's between urban, peri-urban and rural schools combined with exploratory talk and suggest 'social capital', local environments and initial language capability as possible intervening factors. Further research is needed in order to confirm the influence of these variables.

Turn-taking

In the pilot study we saw no real change regarding quantitative symmetry and interactive dominance and recession, at least not in the non-curricular discussion. We did find substantial change when we analysed the problem solving discussion: after the intervention interactive dominance and recession had vanished, while in two of the three conversations turn-taking had become more symmetrical.

In the main study turn-taking became more democratic during the non-curricular discussion in the target group, though not for all triads. Interactive dominance/recession did not change either, though it involved only few triads. No changes were seen in the control group. In the problem solving discussion the target group showed more quantitative symmetry after the intervention, and again there was no change in the control group. It appears that the intervention has caused some positive effects. As Mercer and Littleton (2007) stipulate, without a training in exploratory talk pupils tend to use disputational and/or cumulative talk. Characteristics of disputational talk are that pupils do not really listen to one another and interrupt each other, which results in more quantitative asymmetry. Our data confirm this postulate only partially, but then perhaps the 12% margin we used to determine symmetry was too small. This may also explain the remarkable difference at the pre-test for the problem discussion between the control and target group.

A similar evolution was observed by Littleton et al. (2005) in whose study initially less democratic group conversations became more democratic or symmetrical after the intervention. We also find confirmation of our study in Rajala et al. (2012) when we say that individual pupils can only master exploratory talk if they can participate in group exploratory talk. As democratic participation in conversation increases, so can the use of exploratory talk.

Long utterances

Compared to the pre-tests the pilot group enlarged its amount of long utterances considerably, one triad taking it up to 23 in their ten minute conversation. Progress was most visible in the non-curricular discussion. In the problem solving discussion one triad evolved from 1 to 13 long utterances, while the other two ended up with a higher amount of 4 and 5 long utterances. The intervention clearly showed positive effects.

In the main study, the intervention had a positive effect on the target groups, especially during the problem solving discussion (increase from roughly 12 to 16 arguments on average). Progress was less outspoken in the non-curricular discussions, but still positive. Simultaneously, there was no progress in the control groups. We conclude that target pupils have learnt to elaborate more while control pupils have not. This is confirmed by Mercer et al. (1999) and Wegerif (2005), who saw a positive correlation between the number of long utterances, the number of key words and the number of correct answers to the problem solving test (see also Sutherland (2006).

Arguments

On the whole, the intervention has stimulated the pupils of the pilot group to use more arguments and to be more explicit about them. The pupils' quality of those arguments has also increased, especially during the non-curricular discussion. This is not surprising, as during this discussion the pupils were not working with learning materials they could (non-verbally) refer to. When solving Raven's Progressive Matrices, pupils were inclined to talk more implicitly as they 'replaced' utterances by pointing at the puzzles and used the test sheets to demonstrate what they meant.

The main study showed similar positive results for the non-curricular discussion. Nothing much changed in the control groups, but in the target groups the average number of arguments increased. All triads of the target groups made progress, though not to the same extent. In the problem solving discussion the average number of arguments was considerably higher than in the non-curricular discussion. This was found in both tests. This not illogical, as the number of issues to discuss and solve in Raven's was much larger than the number of discussion topics in the non-curricular discussion.

As we hypothesised, the quality of arguments during the non-curricular discussion did not augment in the control groups, while it did in the target groups. Results are in line with the scores for the quantity of arguments, but again not all triads started and evolved at the same rate and pace. Like in the pilot study we noticed that the quality of arguments was higher during the non-curricular discussion than during the problem solving discussion: while explicit and semi-explicit arguments were more dominant in the non-curricular discussions, rudimentary and implicit arguments dominated the problem solving discussions. Again the explanation is to be found in the presence respectively the absence of lesson materials. Whenever materials are used, like in the problem solving discussion, pupils' arguments show more signalling words and deixis, generating more implicit arguments. Nevertheless, our figures show that even then, the target group triads raised not only the quantity but also the quality of their arguments while doing Raven's, whereas the control group did not.

In short, the intervention has stimulated target triads to use more arguments and to increase their quality. These findings are confirmed by Rojas-Drummond et al. (2006). Especially after solving Raven's as a post-test the increased use of arguments was quite noticeable in their study among 11- and 12-year-old children in Mexican primary schools: 'The experimental group produced a total of 42 arguments in the pre-test and 106 in the post-test (almost triple). Furthermore, a more thorough analysis of these latter arguments revealed that they not only increased in quantity but also in quality: they were more clear, coherent, explicit, precise and concise in the post-test' (Rojas-Drummond et al., 2006, p. 88). In an earlier study by Rojas-

Drummond and Zapata (2004), very similar to ours, experimental triads of 11- and 12-year-old pupils who learnt to master exploratory talk also produced significantly more and better arguments than the control triads.

RQ 4. What effects does the use of exploratory talk have on pupils' problem solving skills at group and at individual level?

Based on a number of experiments conducted in the UK and Mexico, the design of which was broadly replicated in this study, Wegerif et al. (2005) concluded that the Thinking Together Approach[31] 'reliably leads to gains on reasoning tests of between 5% and 10% for individuals and between 10% and 15% for groups' (Wegerif et al., 2005, p. 43). The results of this study, though not replicating the Thinking Together Approach entirely, confirm the effects of exploratory talk on reasoning skills in a similar way.

Effects at group level

As mentioned before, we used Raven's Standard Progressive Matrices to measure the pupils' problem solving skills at group level. For that purpose we split up the test into two equal parts, respecting its logical build-up. We used one part for the pre-test and the other for the post-test. The pilot study appeared indicative for the results of the main study. The pilot class scored significantly higher on Raven's at the post-test compared to the pre-test. On average, the three triads improved their scores with 30.6%.

In the main study, the target group increased its score after the intervention significantly with an average of nearly 13%. The control groups score also improved but the difference with the pre-test scores was not significant. We conclude that the intervention had a positive significant effect on the target groups and our results are confirmed by a number of other studies (Fernandez et al., 2001; Mercer et al., 1999; Rojas-Drummond et al., 2006; Rojas-Drummond & Mercer, 2003; Rojas-Drummond & Zapata, 2004; Wegerif, 1996b; Wegerif, 2005; Wegerif & Mercer, 1997a).

Effects at individual level

In order to measure these effects we used an adapted version of Raven's Coloured Progressive Matrices. Following Vygotsky, we expected pupils to not only improve their problem solving skills at group level but also individually, as they would implicitly internalise the group

[31] A dialogue-based approach to the development of children's thinking and learning, developed at the University of Cambridge, UK, which includes the learning of exploratory talk. Cf. https://thinkingtogether.educ.cam.ac.uk

strategies and turn these into individual learning strategies. This was confirmed convincingly in the pilot class and partly in the main study.

In the pilot group individual scores increased significantly with 10.5%, confirming results of similar studies, e.g. Mercer (1999). In the main study we saw less individual progress, i.e. 4.8%. In the control group similar progress was found, but pupils of the target group improved their problem solving skills significantly, whereas the pupils of the control group did not. Here, school differences in the way the experiment was implemented must be taken into account. In their study Wegerif, Mercer and Dawes (1999) noticed that the intervention programme similar to this study had been carried out more carefully and comprehensively in schools A1 and B1 than in school C1. 'Different teachers modelled the use of exploratory talk differently' (Wegerif et al., 1999, p. 499). Accordingly, group scores for Raven's increased by over 10% in the best performing schools (A1 and B1). High standard deviations make us suspect that similar mechanisms have had an influence on our experiment. Some teachers, for instance, had considerably more experience organising collaborative activities than others. Apart from this, after finding significant gains in problem solving skills of target group children at the individual level, Rojas-Drummond et al. (2003) admitted that the individual effect is difficult to achieve, as in a previous study they had only found effects at group level.

It is noticeable that the average individual starting score for pupils of the target groups was higher than that of the pupils of the control groups. We can partly explain this by pointing at – rather coincidental – individual differences, an explanation which we also find in Mercer et al. (1999). Based on an experiment with 60 British children, aged nine and ten, the authors noted that not all pupils – either individual or in group – showed progress solving Raven's, even if their gains in exploratory talk were striking. The high standard deviations in our analyses suggest the same phenomenon. Mercer et al. (1999) found it difficult to explain this and referred to situational factors which are not uncommon in educational research, and to the absence of children at key points of data collection. To this we wish to add our observations that some triads already used many key words in an exploratory context from the beginning, while others did not. Some triads improved their quality of argumentation but not necessarily increased their use of key words in context (see also Herrlitz-Biro, 2010). This makes it more difficult to prove correlations between key words use and group scores on Raven's. Nevertheless, we tend to agree with Mercer et al.'s (1999) suggestions that a 'larger-scale implementation could have provided better data for statistical analysis' (Mercer et al., 1999, p. 109). Taking all this into account and looking at our data, though, we feel safe enough to conclude that learning strategies which were mastered at group level also had a – moderate – effect at individual level. Pupils have not only increased their problem solving skills collectively but also individually.

RQ 5. Which variables explain evolutions in the individual use of key words for exploratory talk?

We took four independent variables into consideration which might influence the individual use of exploratory talk, i.e. key words in context: gender, GOK indicators[32], spelling and mathematics skills.

Gender

As exploratory talk focuses away from identity onto the quality of shared reasoning (Wegerif et al., 1997a), the (gender-)equalising effect inherent to the application of the ground rules may very well show, especially in those cultures where gender differences tend to be more emphasised. In such cultures boys' gender identity may amplify the perception that they should take the leading role during collaborative activities (Wegerif et al., 2005). This would certainly show in disputational and to a lesser extent in cumulative talk, where the formation of a group identity is far from obvious, as social identities are considered more important than the task or than shared reasoning (see also Polo et al., 2015).

Our data did not reveal any significant girl-boy differences as far as progress in key word use is concerned, but we did notice something interesting. On average, in the control group, girls used 14 key words in context less than boys after the intervention in the non-curricular discussion, but in the target groups, girls increased their number of key words in context by 35 compared with the control groups and made 14 key words more progress than boys. In the problem-solving discussion progress of the target group boys made slightly more progress than girls. As no significant differences in progress were found, any generalisations are out of place, but it would be interesting to investigate whether 'problem solving' induces boys to use key words for exploratory talk at individual level more than girls while a more divergent assignment makes girls use more key words.

Other studies did not reveal strong gender differences, either. In his study on the effect of goals instructions on learners' reasoning and argumentation Nussbaum (2005) found no signicant main or interactive effects of gender. Topping and Trickey (2007a) investigated the effects of Philosophy For Children (P4C) which was implemented for one hour per week across a 16-month period on 10- to 12-year-old pupils. The results were that pupils who participated in P4C showed significant standardised gains in verbal and non-verbal reasoning ability, but this was largely irrespective of gender. The near absence of significant gender influences in our

[32] GOK indicatoren or 'indicatoren gelijke onderwijskansen' (equal opportunities for education) were introduced in Flemish education in 2001 as indicators for extra individual support. Schools receive additional financial means to organise education for pupils who meet criteria like low economic family status, limited language skills, etc. In the course of this study GOK was renamed SES (socioeconomic status), sharing the same goals. In this study we consistently refer to GOK.

study may be seen as a confirmation for this effect (at least the ground rules did not create any gender differences, which would be very surprising anyway).

GOK indicators

As in the pilot group all pupils were assigned one or more GOK indicators, it is not possible to draw any conclusions about the influence of this variable for this group. What we do know is that the pilot group made significant progress in Raven's and used significantly more key words after the intervention, both at group and individual level. Wegerif, Mercer and Dawes (1999) hypothesise that middle-class children may be more familiar with the 'ground rules' of exploratory talk than those from the lower-income families (like our pilot group), which would explain this group's collective and individual progress.

Statistical analyses of this variable in the main study, where 20 pupils (control group, n = 9; target group, n = 11) had GOK indicators, did not reveal any significant impact of this variable either. Perhaps, a population of 20 pupils, let alone 9 or 11, was too small to do proper statistics and generalise findings. Further, it is important to note that the data do not specify whether individual pupils were given GOK indication based on one parameter or more. This means that for an individual pupil, GOK indication based on only one parameter may have a smaller impact on his school performance than when more parameters are involved. Nevertheless, we want to discuss some of our findings in more detail.

We noticed that eight GOK pupils of the target group improved their score for Raven's post-test, while three pupils scored equal to the pre-test. It is interesting to see that no GOK pupil had a lower score for Raven's after the intervention. It is conceivable, of course, that these pupils experienced the full benefit of working in groups. As all GOK pupils were distributed equally over the triads and no triad consisted of more than one GOK pupil, this may also suggest that GOK indicators of individual pupils do not have a negative impact on group performance. In the individual test eight GOK pupils improved their score by 10 to 50%, while two pupils scored lower. Wegerif, Mercer, and Dawes (1999) also saw individuals who scored very low on the pre-intervention test improve substantially after learning exploratory talk. They suggest that the intervention taught these pupils strategies for problem solving which middle-class pupils already possessed (Wegerif, Mercer & Dawes, 1999). Comparatively, we noticed that individual GOK pupils made more 'key word progress' compared to non-GOK pupils after the intervention, but then again, this progress was not significant. We lack sufficient data to draw substantial conclusions about this variable. As the socio-economic status of pupils has showed to have significant impact on school performance in international assessment studies such as PISA (OECD, 2016), further research is recommended.

Spelling

As Flemish primary school uses the same valid testing tool to measure spelling skills, we decided to include these data in our analysis. It must be noted, though, that spelling is only one aspect of language skills. It does not say anything about oral skills and not even about certain writing skills. Also, it does not say much about strategic thinking of the kind pupils use when solving a problem, because spelling skills are mostly memory-driven (Sandra, Frisson & Daems, 1999). Nevertheless, we found no correlations whatsoever between key word use and individual spelling skills. Also, we have no knowledge of research that did find such correlation.

Mathematics

Mathematics is the only independent variable for which a significant impact was found on the individual progress of key word use, i.e. in the problem solving conversations. Detailed statistical analysis shows that this increase of key words appears in two of the six categories of key words, which each answer to one of the ground rules: questions (category 3) and the formulation of claims and arguments (category 1). This implies that pupils with high mathematics skills significantly increased their mastery of those ground rules (and the related key words) compared to pupils with low mathematics skills.

As our literature study shows, a considerable number of experiments with positive effects regarding the use of exploratory talk involved science and mathematics as school subjects. Substantial evidence for this specific combination has been delivered by e.g. Murphy (2015), whose literature reviews showed that research in mathematics education has been focusing on communication and interaction in the mathematics classroom for the last two decades, and by Topping and Trickey (2014) and Rabel and Wooldridge (2013), who stress the importance of dialogic teaching and exploratory talk as a peer learning tool in mathematics classes. A study of peer tutoring in mathematics even showed significant increases in use of math words and strategic dialog (Topping, Campbell, Douglas and Smith, 2003). Several experiments show that the enhancement of problem solving skills through the mastery of exploratory talk often correlates with improvements in maths problem solving tasks (Mercer, 1996; Mercer & Littleton, 2007; Mercer et al., 1999; Rojas-Drummond et al., 2006; Rojas-Drummond & Mercer, 2003; Schmitz & Winskel, 2008; Wegerif, Mercer & Dawes, 1999; Wheeldon, 2006). Indirect proof of improved math skills is also provided by higher scores on Raven's after learning exploratory talk, as the aim of this test is essentially to find underlying mathematical patterns. This way, our study has contributed to the growing consensus that there is a direct relationship between exploratory talk and mathematical (problem solving) thinking. To our knowledge, however, our study is the first to find a relationship between individual mathematical skills and the increased use of exploratory talk in collaborative problem solving tasks, that is, if we look at the use of key words in context.

Conclusion

Summing up, the findings of our empirical study are confirmed by earlier research (Mercer, 1996; Mercer & Sams, 2006; Mercer et al., 1999; Rojas-Drummond & Zapata, 2004; Wegerif, 1996b; Wegerif & Mercer, 1997b). It is clear that exploratory talk can be taught/learned and if pupils make use of it consistently, they use language in such a way that they learn more from one another and improve group reasoning. Direct observed effects are that pupils increase their argumentation skills, work better together as they master the ground rules and increase their problem solving ability at group and individual level. Simultaneously problem solving skills improved at group ànd at individual level, which confirms Vygotsky 's claim that social learning precedes individual learning. This way, exploratory talk indeed reflects an educationally effective intellectual activity, a social mode of thinking (Mercer, 1996, 2004). The fact that the only individual variable to have any influence on the individual development of exploratory talk (i.e. the increased use of key words in context) is mathematics skills, which – moreover – only shows during the problem solving discussion, suggests that improvement in group results after the intervention can only in a limited way be accounted for by individual improvement regarding reasoning skills. This is confirmed by Wegerif, Mercer and Dawes (1999). The reasoning ability involved in the individual problem solving test is mediated to a large extent by social interaction.

Chapter 6: Discussion

We wish to conclude this dissertation by addressing a number of points for discussion. We will address theoretical issues, the definition of exploratory talk, its study in general and some methodological aspects. Further, emerging thoughts will be presented on aspects like respondents and intervening variables and the generalisation of our findings. We will finish this chapter by making suggestions for further research and for the implementation of exploratory talk in Flemish education. Discussion points raised by our literature study will be integrated as indicated in Chapter 3 (sections 3.1.5 and 3.2.5).

About Vygotsky and interthinking

Interventional studies like ours have provided results 'which link sociocultural theory and educational practice' (Rojas-Drummond & Mercer, 2003, p. 110). From a theoretical viewpoint, Vygotsky's claim is supported that learning starts on a social level and proceeds – via intermental activities or 'interthinking' (Littleton & Mercer, 2013; Mercer, 2004) – to the individual level (intramental). During this process language is used as a cultural tool which 'not only enables collective thinking to become more effective but also promotes development of individual reasoning' (Rojas-Drummond & Mercer, 2003, p. 110). This process especially seems to benefit problem solving.

One could argue that there is no such thing as group thinking or 'interthinking', only a well-organised and peer induced line of individual thinking. Strictly speaking, this would imply that collaborative activities do not have an educational added value at group level. If that were so, however, the group results at Raven's would never exceed the highest individual result of that group (Wegerif et al., 2017), while our data show quite the opposite.

The question that remains, is how exploratory talk exactly works. Problem solving, as Wegerif (2008) postulates, seems to rely more on the spontaneous generation of new metaphors as 'ways of seeing' than upon explicit reasoning. This idea is very similar to the so-called Aha-experience or the sudden appearance of a solution through insight (Topolinski & Reber, 2010) which finds its roots in the typically human impulse for pattern recognition as was postulated by Gestalt psychologists (Köhler, 1970). It is not our intention to challenge these claims, but if the ground rules prepare the ground for (sudden) meaning giving at group level and hold in fact the key mechanism for problem solving, then exploratory talk must be the visible, hearable and tangible engine that drives the process of insight building. If one pupil finds a solution to a puzzle in Raven's, exploratory talk makes him explain this insight and justify it. By doing so, he may induce the Aha-experience of his peers, who either confirm or question his insight, for even an Aha-experience can lead an individual to wrong conclusions. This is the reason why group scores exceed individual scores. This makes language indeed a tool for reasoning and learning, just like oars can be very effective tools for rowing a boat, if used collectively and in a synchronous way. This way, Mercer's (2000) Intermental Development Zone becomes what

Wegerif (2005) has called the (shared) ‘space of reflection’. Both are dialogical phenomena. As a result the Zone of Proximal Development, which Vygotsky first presented as an individual characteristic, can also be applied to group processes. Further, the ZPD is expanded by exploratory talk whereas disputational talk rather seems to narrow it down and cumulative talk leaves it unaltered (Fernandez et al., 2001). The way this happens is very recogniseable to every teacher: pupils unconsciously take on a scaffolding (but not a tutorial) role and use language to support each other’s learning. Interthinking then also works on an attitudinal level: pupils not only improve their reasoning skills, they also improve their social skills, as group identity includes the learning and refining of social skills. Though every pupil has to acquire these individually, with reflection and inner speech guiding them, again this would not be possible outside a context of collective learning and the formation of group identity.

It is probably only a matter of time before neuroscience adds more answers to the above. So far, it is virtually impossible to put a group of pupils under a scanner to observe what happens in their brains during collaborative activities, but Van Camp et al. (2015) report indirect indications that interactive learning actually shows in the limbic system of the human brain. Babies older than nine months only learn a new language when they hear the words spoken by a real person and not when the voice of that same person is played on audio or video. And gamblers do not only decide ‘which card to play’ based on their own experiences but also on the way other gamblers have reacted on the choices they made earlier. To Van Camp et al. (2015) the first example demonstrates interactive learning, while the second adds to this that learning happens thanks to a shared effect of previous experiences. Areas in the brain which are activated during collaborative activities appear to be the same areas which are activated when we receive a reward or a punishment and when we learn something. Last but not least, Van Camp et al. (2015) pinpoint that collaborative strategies will generate less learning effects when applied in an individualistic and competitive school culture, which is exactly what exploratory talk can compensate for. Clearly, further research must cast more light on this matter.

About defining exploratory talk

When people start a discussion they will most likely voice their opinion, make personal claims and defend these. Some will do this rationally, some will quickly bring in emotion, most will do both. Most will also sin against one or more of the ground rules for exploratory talk, as is most prominently illustrated by disputational talk. The basic idea behind disputational talk is ‘I want to win this’. Therefore disputational talk is by definition more competitive than cumulative and exploratory talk. Debates are good examples of this, though the degree of competitiveness may still vary.

When people engage in a brainstorm, they will most likely voice their opinion, sum up ideas and perhaps voice arguments to support their claim. Brainstorms are a smoother type of talk than a debate, as nobody gains profit from wanting to ‘win the conversation’. Only as soon as it comes to choosing the best idea from all that has come up cumulative talk may become exploratory or even disputational. By restricting to cumulative talk participants miss opportunities to make the best choice, find the best idea, draw the right conclusion.

So, what kind of talk would a problem solving discussion be? In order to solve a problem collectively it is necessary that participants voice their opinion, gather ideas, defend claims and come to an agreement. As the problem usually lies outside the individual, the basic idea behind the discussion would be: ‘This is *our* problem and *we* have to solve it’, making emotional interventions and negative influences like personal agendas less likely. Problem solving discussions therefore offer participants good chances and good reasons to embrace the ground rules of exploratory talk. This does not mean that problem solving discussions cannot be cumulative or disputational, but then the result – solving the problem – will often be less successful.

This study has confirmed the findings of many researchers that exploratory talk is more rational and has more educational potential than cumulative and disputational talk, and that it is probably the best type of talk when it comes to solving problems in groups. Based on recent literature it may be obvious, though, that the triad typology (cumulative – disputational – exploratory) is being tested. This brings us to suggestions that were made by other researchers.

First, we accept the idea that exploratory talk is the most valuable kind of talk for learning, but find it inappropriate to speak of ‘lowest’ and ‘lower level’ of talk as far as disputational and cumulative talk are concerned, as in Schmitz and Winskel (2008). Cumulative talk can be very valuable during e.g. brainstorms or any occasion in which the abundant gathering of (creative) ideas is important. Disputational talk is what often characterise s persuasive communication, e.g. in debates. Depending on the targets set by the teacher (and the curriculum), both types of talk can be very relevant in an educational context. A hierarchical structure would also suggest that the boundaries between the three modes of thinking are clear, which is rarely the case.

Second, we can only partly agree with Rojas-Drummond and Zapata (2004) that the triad disputational-cumulative-exploratory may be too artificial as a taxonomy: they argue that the contrast between exploratory and cumulative talk is too artificial, as it depends on how one defines ‘explicit reasoning’. Explicit reasoning indeed asks for further defining. When pupils start working on a group task, their use of language is likely to differ when they are provided with ‘materials’ or not. In a class setting where pupils had to work together solving problems presented by a computer program, Wegerif et al. (2003) already suggested that computer enabled reasoning may be more implicit than when students rely on words alone. Also, in their qualitative analyses of classroom talk, Herrlitz-Biro et al. (2013) found that pupils can effectively use explorative talk without there being many linguistic features accounting for that talk. But this does not mean that the taxonomy is flawed.

Third, we have certain doubts about the addition of *operational talk* as a fourth type of talk, which Nikolaidou (2012) defines as ‘peers’ utterances relating to operational transactions with regard to talk and software respectively' (Nikolaidou, 2012, p. 744). First, we believe this concept need more refining before one can speak of an extra type within a taxonomy. As it is described now, operational talk seems restricted to talk-about-talk and talk about handling hard/software. Especially as far as the latter is concerned, we think speaking of another type of talk seems rather task or context dependent. We suggest that operational talk may be detectable during any kind of group activities in which pupils have to handle lesson materials, and perhaps different lesson materials will generate different (operational) talk, cf. Tin's (2003) study on divergent vs. convergent tasks. Further, we do not see any reason to consider operational talk as a conceptual variant of talk on the same level as cumulative, disputational and exploratory talk, because of the absence of Polo et al.'s (2015) notion of (self-)identity. Talking about 'which button to press', is not a matter of conflict of people nor ideas, and if it is, talk becomes either cumulative, disputational or exploratory. In this respect, it remains unclear whether operational utterances may not be an intrinsic part of exploratory talk, as they may express a change of attitude due to the knowledge and application of the ground rules of exploratory talk (Dawes et al., 2010). Finally, if operational talk were to be a fourth type of talk, we see no reason why other types should not be added, e.g. *divertive talk* or what Wegerif (2008) calls *playful talk* for those instances during which pupils leave their 'learning context' to make jokes or tell stories. But to what extent can we then uphold the notion of social modes of thinking (Mercer, 1996)?

Fourth, Yaguchi et al. (2010) explicit reference to 'a more affective attitude toward the listener' as a characteristic of exploratory talk may be a valuable addition to Mercer's definition, as it fortifies and clarifies the notion of joint consideration of ideas. Nikolaidu (2012) also adds *reflective talk* as an extra type, defining it as: 'Peers engage critically and constructively express a self-reflective thinking.' However, parallels with other studies (Bowskill, 2010; Chick, 2015; Mayher, 1990, and Riley, 2006) and with the notion of *individual elaboration* in Herrlitz-Biro et al. (2013) suggest that self-reflective thinking may not be a separate type of talk on the same level as exploratory talk, as Nikolaidou (2012) seems to propose by clustering both types, but rather a type of talk which is induced by and may even be a subtype of exploratory talk. Exploring thought in your own mind relates very much to Vermunt's (1993) theories on metacognition for learning. In that context, reflective talk may be considered as another mode of thinking indeed, somewhat like Vygotsky's inner speech, where reflection is in fact the explicitation of self-exploratory talk. Clustering certain speech acts and the key words they induce may very well lead to the formulation of other types of talk, but not necessarily of another *social* mode of thinking, though. Yet, the idea is very much worth pursuing.

Finally, the emerging discussion whether the notion of exploratory talk should be changed into the more Bakhtinian notion of *dialogic talk* (Brown, 2016; Wegerif, 2013) or even into Dourneen's (2013) *dialogic reason*, seems a philosophical rather than a pragmatic issue. In *The Centrality of Exploratory Talk in Dialogic Teaching and Learning,* McConaghy (2014) joins exploratory talk as a tool of inquiry with dialogic teaching in the classroom (see also Alexander, 2010) which she calls 'a pedagogical practice'. In this respect, exploratory talk and dialogic teaching are interdependent and interrelated (McConaghy, 2014). One could argue that speaking of dialogic talk instead of exploratory talk may symbolise this joining of notions. We do not want to join a philosophical discussion about this issue, as this dissertation focuses on a more pragmatic approach on learning, for which we find the notion of exploratory talk more transparent to work with than dialogic talk.

About studying exploratory talk

Study 1 has taught us that most research does not really measure whether a conversation is exploratory or not, but only to what extent it is exploratory, either in itself or compared to previous conversations (e.g. the use of key words in context in a post-test vs. a pre-test discussion). Is it possible to establish such an 'exploratory checkpoint'? Only two studies went on to explore this path.

Polo et al. (2015) operationalised Mercer and Wegerif's concepts into five indicators of group talk: (1) whether assertions and refutations are justified, (2) whether the students elaborate on the argumentative content of previous turns, (3) whether they critically evaluate each other's arguments, (4) whether they take everybody into account when making the decision for the collective vote, (5) whether a particular students' talk during the following large group debate integrated the rest of the group's supporting or opposing argumentation or only voice his own initial ideas (Polo et al., 2015). When analysing talk in three case studies they considered exploratory talk to be achieved when the five indicators were positive, but at the same they asked themselves the question: what happens when some but not all the indicators are positive? They concluded that 'students' footing is not necessarily the same during a whole interaction and such a balance can be modified according to local interactional aims. For instance, subsequences of disputational talk expressing the different competing views seem sometimes […] to be normal openings for exploratory sequences, and must not be given derogatory status' (Polo et al., 2015, p. 28). Polo et al. agree with Mercer (1995) that the use of ground rules or other kinds of preparation (see Janssen et al., 2010) is crucial in order to implement the use of exploratory talk, but they warn for over-scripting (e.g. by giving participants in a group a maximum number of turns). In addition, they found that preparation tools like ground rules do not entirely prevent pupils to use cumulative and disputational talk from group discussions, as these may serve specific functions of wider, globally exploratory sequence. In some cases disputational talk may even trigger initiations of exploratory talk (Rajala et al., 2012). Finally, pupils may need time to 'roll into' the exploratory discourse pattern.

In our opinion, there are three ways to look at the issue of an 'exploratory checkpoint'. From a linguistic point of view, trying to determine whether conversation is exploratory or not, can only be done by quantifying the language items used (words and phrases). From a communicative point of view, quantifying certain speech acts (asking for explanation, persuading someone) would be necessary. From a functional view, lastly, one would have to quantify the outcomes of the conversation (new insights, problems solved, agreements on topics …). In all those cases one would have to analyse a significant number of conversations held in a number of different contexts (problem solving discussion, moral debates, brainstorms, etc) and based on these analyses determine a cut-off point. If no other variables were involved, we believe this might be possible, but one only has to consider the participants of a conversation to realise there are plenty of intervening variables, like age, gender, ability, language skills, motivation, etc. This is what Perry (1999) suggests when he says that the point of view learners approach knowledge from depends on their personal context of learning. Besides, even if a conversation would be labelled exploratory on one ground, it may not be on the other. A good example of this are the analyses of Herrlitz-Biro et al. (2013) who found that exploratory talk can sometimes lack the linguistic markers that define it.

In another study, dealing with argument development among undergraduate students Golanics and Nussbaum (2008) analysed talk in triads of pupils and marked the dimension of exploratory talk on a 5-point scale. They assigned marks as following: 5 if all three group members are exploratory (critical but flexible, willing to concede, etc); 4 if two or three are exploratory; 3 if one of the three is exploratory; 2 if group members either were all cumulative or disputational; 1 if the group members mostly repeated one another's comments. After analysis the inter-rater agreement was 77% on exploratory discourse. It is clear that Golanics and Nussbaum (2008) have very much narrowed down the analytical focus of exploratory talk, leaving out e.g. the matter of key words, turn-taking, length of utterances, etc. By using a 5-point marking scale both researchers also seem to agree there is no real cut-off point to call a conversation exploratory, only as much as a continuum of 'less or more' of such talk. In short, their scale does not indicate that a conversation is exploratory from start to finish either. And even if a conversation would show ample proof of the occurrence of pre-set indicators, short shifts to cumulative or even disputational talk may still occur, as said before. From a technical point of view it may be justified to assess the quality of such a – only partly exploratory – conversation as negative. However, when looking at the content of a conversation, i.e. the lines of thought, the expression of opinions, the construction of arguments, the agreements reached, etc, the quality of the conversation may be excellent, despite the lack of consistent exploratory discourse.

More interestingly, even though they express the thought that 'by its nature, exploratory discourse characterise s the way individuals in a group interact, and so is only meaningful at the group level' (Golanics & Nussbaum, 2008, p. 175), the fact remains they analysed the extent to which exploratory talk was visible on an individual level. This, together with Perry's (1999) claims about individual context-driven learning, is precisely the reason why we decided to look at individual variables which might influence progress in the use of exploratory talk.

About some methodological issues

In this section we address some more general and some more specific issues concerning our research design. We will discuss switch replication, the relationship between our pilot and main study, research instruments and the indicators for exploratory talk.

Switch replication

For practical and organisational reasons we applied switch replication in school 2, which means that the two control groups of this group became target groups after the control phase. Expressing some self-criticism we believe this may have added an unwanted variable to our intervention: changes in group atmosphere. The results of the problem solving post-test was positive but in a less outspoken way than in the other schools, where switch replication was not applied.

There are two ways to look at this. First, there is the human factor: control ànd target group triads had to work together for nearly a whole school year, which proved very demanding for some pupils. During the intervention we noticed that in some triads cooperation became more difficult and less efficient as the pupils grew bored of one another. In one triad small quarrels popped up: indicators of exploratory talk we had found in the control condition evaporated to be replaced by disputational utterances. Kagan (2014) proposes keeping groups of pupils together for at least (but not much longer than) five weeks in order to develop an optimal working relationship and Mercer (1995) suggests that friendly working relationships are a benefit for the development of exploratory talk. But if quarrels and an unfriendly atmosphere take over, the reverse is quite conceivable. Unwillingness to engage in (the learning of) exploratory talk works contraproductively on the process (Robins, 2011). We believe swapping groups would have helped pupils to develop new working relationships and benefited the experiment, especially in those groups where the atmosphere had turned sour.

There may, however, be another explanation why results in this school were less than expected. Wegerif, Mercer, and Dawes (1999) describe the failure of a similar intervention programme in one of their target schools. They hypothesise that training children to use exploratory talk 'has the most positive effect on children's joint attempt to solve the Raven's problems when the children involved are from lower socio-economic groups' (Wegerif, Mercer & Dawes, 1999, p. 513). Wegerif suggests that middle-class children might have been more familiar with the ground rules of exploratory talk. One of his arguments is that some individual pupils who scored very low on the pre-test improved substantially after coaching in exploratory talk. Their initial failure 'was due to the lack of quite simple strategies which the intervention course was able to provide them with.'(Wegerif, Mercer & Dawes, 1999, p. 513). Though in our study the target group with most children from lower socio-economic groups was the pilot group, resemblance is striking: some individual pupils of that group scored very low on Raven's in the pre-test and indeed made strong progress after the intervention. In school 2, a suburb school with middle- and high-class children which had as good as no pupils with GOK indicators, progress of problem solving skills was poorest. Considering yet another viewpoint, some triads may have developed so much collaborative routines during the control phase that further – exploratory – progress became less probable or perhaps tangible. All this may have influenced the statistics we have presented in this dissertation, but comparison with the mastery of exploratory talk at the same time fortifies our assumption that solving problems and talking exploratively are interrelated. We suggest that mastering exploratory talk offers pupils cognitive strategies which make them approach problem solving assignments with more eagerness, as they are perceived as interesting challenges.

Concluding, we do not believe switch replication to be a bad idea, as other factors may have played a role (the results in school 4 for Raven's were also less), but if we were to replicate our own study, we would swap triads after the control phase. Another argument for this is that when a school year is halfway through, the class teacher will know very well which pupils form good triads for collaborative tasks.

Pilot and main study

Especially in our pilot study, video recordings showed many conversations that were either disputational or cumulative, or both: pupils used a lot of non-verbal language to make their claim, interrupted each other, looked away from the pupil who was talking and did not seem to listen to one another. The lack of verbal skills of some pupils sometimes seemed to enhance their non-verbal behaviour, leaving few possibilities for exploratory talk to occur spontaneously. Pre-test conversations in the main study also showed many features of cumulative and/or disputational talk. The video recordings showed better cooperation and pupils seemed more at ease, but then again, these pupils were older than the pilot school pupils and had better general language skills.

In our conclusions of the pilot study we said that these pupils, many of whom had weak language skills, picked up a number of key words as if they were new vocabulary. The number of key words they used in the post-tests was also considerably higher than in the pre-test and higher than in the main study. We believe this mechanism made them sometimes use these key words the wrong way. Utterances as 'I agree with no. 5' while solving the Raven's test, instead of 'I agree with Meryem[33]' are illustrative: the pupils wrongly assumed that agreeing with something means that you think something is right, while in fact you can only agree with a claim made by a person. This kind of misinterpretation was only found in the pilot school. We hypothesise that children with weak language skills acquire the key words in a different way than children with normal or strong language skills. Reflecting on several of his empirical studies Mercer (2010b) noted that significant progress in mastering exploratory talk was seen

[33] This name is fictitious.

in schools in deprived areas, with lots of pupils from a problematic socioeconomic background. We see parallels with our pilot school and believe it would be very interesting to focus research on those groups of children.

Research instruments: the non-curricular and problem solving discussion

This brings us to the issue of the non-curricular discussion we used in the pre- and post-test of this study. Our analysis shows that the non-curricular discussion differs fundamentally from a problem solving discussion, if only because in the former it was not necessary to come to an agreement. Considering the characteristics of a discussion, as outlined in the previous paragraph, the chances that a discussion is or becomes exploratory seem less than when problem solving or a convergent assignment is at stake. We noticed this in our analysis of key words in context: 'exploratory progress' was less visible in the non-curricular discussion than in the problem solving discussion. Herrlitz-Biro et al. (2013) even integrated problem solving in their (verbal) pre-test: pupils were first given a dilemma about which they first had to form their opinion individually and write their choice down with supporting reasons. After that, they had to discuss their opinions with the other pupils and make a group decision by reaching agreement.

When problem solving is at the heart of the exploratory matter, it seems only logical to focus on those kinds of discussions, also and perhaps especially in testing conditions. This may cast some doubt over the use of a non-curricular discussion whatsoever, but at the same time it sheds light on the use of ground rules and exploratory talk outside the problem solving field. The question that remains is whether the whole experiment, which focused almost exclusively on problem solving tasks, would not better be re-designed and systematically include other types of discussions as well. Teachers would then have to diversify reflections to make pupils understand which ground rules work best in what context and this may in turn add to the use of exploratory talk in various discussions. But this would undoubtedly make studies like this more complex than they are already.

Research instruments: Raven's Progressive Matrices

Many empirical studies we analysed in our literature study use science or maths as their working ground. The rationale behind this is that problem solving skills mostly occur within subjects that include a lot of problem solving, like science and maths. Though we agree that science education – in accordance with constructivist thinking – deserves and even requires more problem solving starting from a familiar context, instead of being taught in a traditionally abstract way, we regret the implication that problem solving should be restricted to science and maths. We strongly believe that problem solving is part of a cross-curricular didactic approach. In other words, it can easily be embedded in subjects such as languages (e.g. grammar), world orientation and religious studies/ethics. Further, we have noticed that in 17 studies problem solving skills are measured using Raven's (Coloured) Progressive Matrices as testing tool, in correlation with evolutions in the use of exploratory talk. The Raven's test is non-cultural and non-verbal. Therefore, it may have more affinity with abstract subject matter and classroom activities as in science and maths, which explains their prevalence.

Yet, some doubts remain about Raven's. We do not question the validity of the Raven's test in itself, but we do question its validity for this and perhaps other studies that used the test for similar purposes. The test consists of 5 'levels' of progressing difficulty, each counting 12 puzzles, where A is the easiest set of puzzles and E the most difficult. Splitting the test in two parts of equal difficulty (one for the pre-test and one for the post-test) results in 5 levels which each count 6 puzzles. Our analyses show that pupils of the target groups made most progress on the D and E level puzzles. Scores are almost perfect for the A and B level and very high for the C level, in the pre- as well as in the post-test. This means that progress for problem solving skills is factually determined on a relatively small basis: 6 D and 6 E puzzles. Despite significant differences between the target and control groups' scores we may wonder if 12 distinguishing puzzles allow for drawing generalising conclusions at group level. On an individual level, however, we did not encounter this phenomenon to the same extent. Individual pupils showed more score variety on all levels (though the A-level still appeared to be very easy).

It must be said that these thoughts already came up during the pilot study, but finding an alternative for Raven's at short notice was not possible. It may be a comfortable thought that studies on exploratory talk that use Raven's very much yield comparable, positive results, but we believed it would be a good idea to develop or search for a problem solving test which serves the exploratory purpose more specifically. It was only shortly after we had finished our experiments that we learned about the newly developed Group Thinking Measure (Wegerif et al., 2017) which combines two tests of equal difficulty by which the correlation between group thinking as well as individual thinking can be measured. At the time when this dissertation was printed, the tool was open for free trial by schools. It will be interesting to learn about the results.

Indicators for exploratory talk

Confirmed by Mercer et al. (1999) our study has shown that the mere count of key words for exploratory talk can only lead to false conclusions about the exploratory nature of talk. First, key words are not a separate language, many are used constantly for other purposes than exploratory talk. Second, and here we find confirmation in the Herrlitz-Biro et al. (2013) study, we have also noticed that pupils sometimes talk exploratively without using such key words. Wegerif (2005) accepts this possibility as well, saying that serious thought can take place implicitly in utterances, even without someone challenging or counter-challenging him. It is one of the reasons why we calculated the use of key words in (exploratory) context and out of (exploratory) context, which showed huge differences between triads. Based solely on the proportional occurrence of key words in context, some triads in this study produced more exploratory talk from the start. As a result they made less statistical progress. This is why we also measured turn-taking, long utterances and arguments. However, this list is not exhaustive. It does not take into account the impact of identity-building during the conversation, non-verbal language and others.

Let us explore this idea a bit more. We hypothesised that exploratory talk would show via an increased use of key words in context, democratic turn-taking, longer utterances and more and better arguments. So far, statistics have been in favour of this hypothesis. However, for all indicators high standard deviations, low minima and high maxima were found. It would be interesting to know whether these four indicators can be ‘summed up’ as a joint indicator of exploratory talk. In other words, do triads that use many key words also produce more and better arguments, is their turn-taking more democratic and do they produce longer utterances altogether after the intervention? We could, for instance, give each triad 1 point for each indicator they have a higher score for after the intervention. This means that when a triad improves on all four indicators, it is given 4 points. When a triad outperforms for three, two, one or no indicators, it receives respectively 3, 2, 1 and 0 points. When scores are equal and in favour of exploratory talk (e.g. when turn-taking is democratic already from the beginning), triads receive half a point. Logic then dictates that target triads will collect more points than control triads. Though outside the scope of this study, we tried this out and found significantly higher scores in the target groups than in the control groups. However, again standard deviation was high in both groups: in the target groups some triads scored only 1 or 2 points, while in the control groups some triads scored 3 or 4 points. This may also indicate that a linear model which simply counts the presence of four indicators does not suffice. Is democratic turn-taking, for instance, equally important as the use of key words in context? And if not, which procedure can help us to adjust the model in order to give turn-taking its proper weight? We believe qualitative analysis is necessary to answer these questions before further calculations and statistical analysis can take place.

From a theoretical point of view we believe it may be interesting to construct a more elaborate and refining conceptual framework for exploratory talk. Such framework would also be beneficial to the one context where exploratory talk has proven its value before any other: that of the classroom. It could be a fine basis to work out effective exploratory didactics and lesson materials for teachers.

Another idea is more pragmatic and came up by coincidence. More or less simultaneous with this study colleagues of mine and two other Flemish universities developed D-PAC or the Digital Platform for the Assessment of Competencies. D-PAC is a web based tool which uses the method of Comparative Judgement to assess Competencies within a variety of contexts (writing ability, self-reflection, problem solving …) and for different purposes (formative evaluation, peer assessment, research, selection...). Comparative Judgement is based on Thurstone’s Law of Comparative Judgement, which says that comparing two things generates a more reliable assessment than having to assign an absolute score to (only) one thing (van Daal, Lesterhuis, Coertjens, Donche & De Maeyer, 2016).

After we had finished our classroom experiments my colleagues asked me to include some of my data in D-PAC as an example of good practice.

Again, this fell outside the scope of this study, but curiosity won the toss and so the following happened: of our video recorded conversations we selected 10 samples of one minute each. Samples were taken from pre- and post-test conversations of 10 different triads. Some extracts we knew were good examples of exploratory talk, others were plainly disputational, cumulative or only partly exploratory. All samples were put onto the D-PAC platform. A group of some 40 teachers was then asked to log into D-PAC individually. Once logged in, they were told they would have to assess ten different conversations, based on the use of the ground rules of exploratory talk, which were visible on their screens via a drop-down menu. The 10 video fragments were shown in randomised pairs and after viewing both, the participants had to indicate which conversation met the ground rules of exploratory talk best. They were also asked to write a comment (strong and weak points) under each video. After that another pair of conversations was shown and the teachers repeated their assessment. As they proceeded, they were not only shown new fragments, but also fragments they had already evaluated. As such they slowly ranked all ten conversations based on the criteria mentioned above.

After all teachers had finished their assessment, D-PAC calculated the average reliability of the rank-order, the result being .84, i.e. high reliability. The comments soon made clear that some teachers had indeed kept an eye on the ground rules while assessing the video fragments, but most had (also) used other criteria, based on their own teaching experiences, like non-verbal behaviour, collaborative attitude and language skills in general. We then compared the rank-order the teachers made with our own rank-order, which was based solely on the occurrence of the four indicators of exploratory talk, and found that it was very similar to that of the 40 teachers. In other words, even without knowing much about exploratory talk, the teachers seemed to recognise and acknowledge its added value for group work and the use of the ground rules. We leave the implications of these findings for what they are but want to make strong recommendations for further research on this topic.

About respondents and intervening variables

In this section we will address gender, classroom culture and teacher behaviour and pedagogic demands for the implementation of exploratory talk

The issue of gender

As we said in the previous chapter, further research is recommended to determine the impact of gender on individual progress in the learning of exploratory talk, but perhaps the research focus should not be the use of key words in context but one of the other indicators of exploratory talk, i.e. turn-taking. In cultures where boys take the lead and girls are discouraged to participate, mixed gender groups would make (a)symmetries of talk and the presence/lack of group identity instantly visible. One striking example of this was found in our pilot study, where during Raven's pre-test one girl, who was part of a completely non-Belgian triad, hardly said a word (she took only 4% of all turns). After the intervention, during the problem solving test, she took 35% of all turns and that same week she improved her personal score on Raven's with almost 50%. Parallels can be drawn with the results of qualitative conversation analysis after teaching exploratory talk to Mexican children, as reported in Wegerif et al. (2005). Similar to our study, mixed triads of girls and boys were formed. In the two boys and one girls groups, girls were encouraged to take on a more leading role in the joint problem solving. One of the hypotheses was that the ground rules would make it easier for girls to speak and for boys to listen to them and to follow their lead. Indeed, the intervention seemed to have caused 'a shift in the role of one girl in the group from being a little subordinated towards taking on a more leading role in the joint program solving' (Wegerif et al., 2005, p.47).

Classroom culture and teacher characteristics

As we analysed the scores for Raven's at individual level, we were first somewhat disappointed that progress of the target groups was not as outspoken as in similar studies. However, high standard deviation – a common feature in our data – made us suspect that the differences between triads might be classroom i.e. school bound. Though the teachers were well briefed before the experiment started, it soon became obvious that school and classroom culture, including the teacher's pedagogic preferences, not only differed but could also have an impact on the experiment. To illustrate this: in school 1 the teacher said group work was one of her favourite work forms, even though for her colleagues it was not. In school 2 the teacher, though being very cooperative during the experiment, said he had little experience with group work. In school 4 the teacher told us she and some of her colleagues very much liked challenging the pupils, asking lots of questions and stimulating them to come up with arguments. It is quite plausible these different viewpoints, experiences and beliefs may have had an effect on what happened in the classroom. In many studies this is a key issue. For instance, Rojas-Drummond (2001) found that modelling was very important while teaching the ground rules of exploratory talk and leading pupils to the process of argumentation that facilitates knowledge construction. Wegerif (1999) found that in one school the intervention programme was implemented by an enthusiastic teacher who was also a researcher and therefore committed to the programme and the study. The transfer to two other schools proved problematic, however, and was called an issue for further research. This sounds quite recognizable as in our study, school 4 teacher had been putting dialogic teaching and 'exploratory mind probing' into practice before the experiment had even started, while school 2 teacher had not. Toppping et al. (2014) shared this experience, finding that some teachers were more able to assimilate what they call ‘the concept of collaborative enquiry’ and put it into practice.

Though the role of the teacher, i.e. the tools, support and scaffolds he provides, may be crucial (Jansen et al., 2010), merely asking teachers to adopt new strategies to engage pupils in exploratory talk is unlikely to be very successful (Harris, 2005; Mercer, 1995). Even when teachers see the added value of knowledge construction through exploratory talk or dialogic teaching, external factors such as a demanding curriculum or a traditional school culture often make them dismiss this as unrealistic (Coultas, 2012; Wegerif, 1996). They may even experience exploratory talk as a contrast to teacher-led conversation in the classroom (Chick, 2015). Some teachers, especially science teachers, also admit they lack the knowhow and the skills to organise opportunities for discussion or do not have the confidence to implement it (Cervetti, 2014; Harris, 2005). How difficult the ‘dialogic shift’ for teachers can be, is described sharply by Barnes (1976): ‘The very presence of a teacher alters the way in which pupils use language, so that they are more likely to be aiming at ‘answers’ which will gain approval than using language to reshape knowledge. Only the most skilful teaching can avoid this.’ (Barnes, 1976, p. 78). It is very important, then, that teachers feel committed to talk to ‘make it work’, that they see the added value of dialogic learning processes and understand how it can fit into their curricula, and that they possess the skills and techniques to put it into practice, either as a participant or as a ‘discourse guide’ who can model exploratory talk and raise his pupils’ metacognitive awareness of how talk can facilitate their collective reasoning and thinking (Webb, 2015; Coultas, 2012; Kerawalla, 2012; Harris, 2005; Maloch, 2002; Wegerif, 1996). Apart from pioneers like Barnes (1976) and Mercer (1995) a considerable number of researchers has laid bare strategies and techniques teachers can use to scaffold group discussions (Hewit, 2014; Gillies, 2013; Wheeldon, 2006; Sutherland, 2005; Rojas-Drummond, 2003; Maloch, 2002). In general, these strategies include the gradual release of responsibility and incorporating more group work in daily classroom practice, modelling, being observant at the start of the process, coaching, making metalinguistic interventions and asking rational thought provoking questions, etc.

The onset for this must be given during teacher education. Student teachers must be given ample opportunities and stimuli to try out dialogic techniques to find out ‘what works’ and to develop a dialogic teaching attitude (Chick, 2015; Topping, 2014; Lofgren, 2013). Chick (2015) suggests that teacher educators should therefore employ exploratory talk during pedagogical discussions as a replacement for the often used educator-led ‘feeding back’, but for this to happen teacher educators need to view learning from a sociocultural perspective. They have to see dialogic interaction as valuable for teaching as well as reflecting on it. But as Fisher (2011) already mentioned, even strong teacher education may not suffice when a trainee’s ‘dialogic willingness’ has been smothered during his own school career. As the author suggests, it takes a dialogic school experience to convince some teachers of the value of dialogic teaching and of exploratory talk (Fisher, 2011).

Pedagogic demands

The younger children are, the more basic lessons are necessary to make them aware of what talking, listening, challenging, agreeing, etc are. As they get older, pupils will be able to master more ground rules and/or be given the opportunity to formulate their own ground rules. It is very important that the implementation of these rules is guarded, at first by the teacher and later by the pupils themselves. For this purpose Mercer and Littleton (2007) and Dawes and Sams (2004) developed various materials which are freely available to teachers and which can be found on the exploratory talk project website Thinking Together. For this study, the basic lessons were adapted to meet the interest of Flemish pupils, but they certainly deserve further refining towards different age groups. Working with talk prompts, for instance, appears to be a very good tool when teaching primary school children, but adolescents are quite likely to dismiss the use of such prompts as childish and irrelevant. For older pupils, the 'guided discovery' approach by which they are stimulated to discover and then use their own set of ground rules in a systematic way, as suggested in Rojas-Drummond et al. (2004), seems more appropriate.

Based on these observations, we see possibilities to work out a progressive learning trajectory of exploratory speech acts and the key words they relate to. For starters, the checklist for problem solving group work turned out to be a helpful tool for the teacher and so did the evaluation forms for the pupils. During the experiment we saw lessons which lacked challenges for pupils to really learn from one another. This reduced group work to working in groups: pupils did not really need each other to get the work done. Didactic guidelines are therefore necessary, not the least for educational publishers.

In that respect Tin (2003) raises the issue which comes first: giving pupils assignments which induce exploratory talk or teaching exploratory talk as a means to successfully finish assignments. As Tin writes: '… what we need to consider here is what causes students to engage in such exploratory talk and whether it is exploratory talk which causes the discovery of truth and expertise or whether it is expertise and knowledge of truth which leads to exploratory talk.' (Tin, 2003, p. 63). We do not believe this issue is at stake as early as in primary education. Tin's experiment involved higher education students who apparently had already mastered exploratory talk to a great extent. Our study (and others) has made clear that this is not the case in primary school. There, pupils first have to learn exploratory talk.

But Tin's study does include a warning: it does not help to try and improve pupils' use of exploratory talk if assignments lack challenge and collaborative needs. Putting it in Vygotskyan terms: assignments must be grounded in the pupils' Zone of Proximal Development. A fortiori, according to Rojas-Drummond et al. (2006) task dependency may very well determine whether explicit reasoning as defined in exploratory talk is required. As we were studying the videos and transcripts of triads who were solving Raven's, we soon noticed that the very easy puzzles were solved without much exploratory talk. Most conversations in that stage of the task were rather cumulative and claims were rarely challenged, as there was no point doing so. So, it seems obvious: for exploratory talk to take place at its fullest extent there must be something to explore, otherwise argumentation risks becoming a means in itself and may even turn into unnatural discourse. It is up to the teacher to see to it that whatever task he provides his pupils with a) has the potential to evoke exploratory talk and b) raises the need to use it.

Again, the question of specification rises: must tasks be problem solving tasks in order to realise the full potential of exploratory talk, and if so, do we not need a more over-arching term for productive discussion in education in order to include exploratory talk which takes place in other contexts (Rojas-Drummond et al., 2006)? In their study, the authors had pupils jointly write a summary of three different texts with related content after discussing, through exploratory talk, the main ideas for such a summary. This test also included some creative thinking. The task was more open-ended and divergent than e.g. Raven's and explicit reasoning appeared to be not so useful for the results to meet the required quality criteria. Because of this, as we saw in Study 1, Rojas-Drummond et al. (2006) propose *co-constructive talk* 'as an inclusive term to characterise the joint efforts of coordination, negotiation and collaboration in various group work activities' (Rojas-Drummond et al., 2006, p. 92). The authors add that further research on this is needed.

We believe an inclusive term may be useful as part of an overall theoretical framework which not only includes exploratory talk but also other types of talk we encountered during our study. At the same time, we wish to warn for a too reductionist view on the use and function of exploratory talk: first, a number of studies on exploratory talk suggest positive outcomes when open-ended, divergent tasks are involved (e.g. Luby, 2014; Robins, 2011; Soter et al., 2009; Nussbaum, 2005). Nevertheless, we need more effect studies involving such tasks if we want to get a grip on contexts where exploratory talk is most fruitful. Second, the question is how convergent the teaching of exploratory talk itself can be. For instance, Rojas-Drummond et al. (2006) mention that in some cases co-construction may be hampered by the need to think exploratively, because pupils believe they cannot express contributions that cannot be justified. It is clear, and our study supports this, that teaching exploratory talk is, in a way, scripted: the ground rules form the scope for learning and, most strongly during the basic lessons, everything is being done for pupils to understand and master these rules. The result of this learning process must be the mastery of a skill but also a strong awareness when it is most appropriate to employ it. When pupils believe they must talk exploratively at all cost during group work, they have a problem with awareness and functional thinking, which is precisely what they need to embed creative thinking in their reasoning processes. In other words, the ground rules are to create 'a space of reflection between participants in which resonance between ideas and images can occur as well as co-construction can occur when participants build creatively on each other's proposals' (Wegerif, 2005). Awareness of which register to use within that space seems to us a matter which deserves further research.

About the generalisation of findings

Nearly all the results of our study, whether positive or not, whether significant or not, show high standard deviations. The differences between triads and – a fortiori – between individual pupils is a common denominator in this study. Averages hide differences between triads. Though in general, the target groups made more progress in acquiring exploratory talk, some control triads also did. Moreover, qualitative analysis in our study showed that some pupils had already acquired exploratory language skills one way or another. Though in most cases children do need systematic instruction to master this type of talk, we cannot ignore the fact that children not always need training to involve themselves in exploratory conversation (see also (Rojas-Drummond & Zapata, 2004).

Partly, this explains why conclusive results were sometimes difficult to establish, because when individual differences are very high, it becomes difficult generalise . By working with five target classes and five control classes we attempted to reach more generalising results than a purely qualitative study would yield. At the same time it is clear that we may have been over-optimistic in pursuing our goals. Clearly, we need a larger study population, a suggestion which was also made by Wegerif, Mercer, and Dawes (1999).

About 'black box' and follow-up research

By posing and answering RQ 5 we have tried to discover more individual contributions to exploratory talk in group work and individual variables that may or may not stimulate this type of talk. Few are other studies that did so. Vygotsky made social learning a preliminary condition to individual learning, but did not theorise how long this process would take and how it evolves. It is fair, however, to assume that internalization is a personal matter: apart from ability, motivation, foreknowledge, etc every learner has their own learning pace. This raises a number of interesting questions: what happens when one pupil starts using key words in context while other group members have not (yet) developed the reflex to do so as well? Does (s)he speed up the learning process of his fellow-pupils? Or does (s)he feel frustrated because the others have fallen behind and stops asking why-questions? And what happens in the brain while exploratory talk is being mastered? Do pupils learn in a straightforward manner or by trial and error? And so on.

Watson's (1920) notion of 'trial and error' leads us to the somewhat behaviourist character of this study and other, similar studies: a group of individuals is given certain stimuli (training in exploratory talk) after which responses (learning effects) are measured. The behaviourist adagium is that learning is essentially a change in behaviour. As behaviourists did not have the means to examine what happened in the human brain while learning is taking place, they called it the 'black box'. And using those words they actually said that the study of learning processes is irrelevant. But after behaviourism came cognitive psychologists who rehabilitated the human brain calling it the 'white box', and started investigating learning processes.

The present study was an effect study, which basically focused on stimulus and response. But as we have some recorded and transcribed conversations in reserve, we believe there is still a lot of ‘white box’ information to be disclosed which now fell outside the scope of this study. This offers possibilities to open the ‘black box’ and see which processes took place between pre- and post-tests. Qualitative as well as quantitative analysis may shed more light on the way mastering exploratory talk evolves, both at group and at individual level, both for every separate indicator of exploratory talk and for each conversation as a whole. It may also explain the huge differences between triads, so that e.g. didactic approaches can be refined. Gender issues may be clarified, as well as the influence of task assignment, mathematics skills, feedback and feed forward reflection, and even - as we also recorded each lesson globally - the aspect of scaffolding by the teacher. Additional analysis of video recordings may provide information about the improvement of the ‘collaborative atmosphere’ and attitudes, intervening aspects such as lesson materials and role shifting in groups. Transfer issues may also be studied, i.e. the way in which exploratory talk shows during various school subjects and the way in which it helps to cross cultural divides.

About exploratory talk in Flemish education

In 1996 Mercer wrote: ‘There is no evidence to suggest that explicit consideration of the ground rules of educational activities has become a normal part of the classroom life in schools or other educational institutions anywhere in the world’ (Mercer, 1996, 375). In the years that followed, it became clear that exploratory talk, as soon as it had proved its educational value in the UK and found its way into the curricula, was there to stay. Gradually, researchers from other countries joined in. In their study within the context of Mexican education, Rojas-Drummond et al. (2004), for instance, found ample reasons to argue that ‘the promotion of argumentative skills should be fundamental as part of the school curriculum’ (Rojas-Drummond et al., 2004, p. 555). Our literature study, Study 1, has made clear that over the last two decades research on the educational value of exploratory talk has become more and more international. It is quite noticeable that teachers who get involved in such research projects, acknowledge this value and so, on the field the message is spreading as well. It is regrettable, then, that more than 40 years after Barnes (1976) launched the concept of exploratory talk, it is still as good as uncharted territory in Flemish (and Dutch) education, both in classroom practice and in research.

We believe it is time for exploration. This is not a call for revolution but for evolution, for the seeds for change are already in our curricula. Compared to, say, 30 years ago, there is more room for oral and written skills in the Flemish language curricula, which leaves opportunities for exploratory talk wide open. Science curricula demand that teachers, before introducing e.g. the Law of Gravity, the process of osmosis or the reason why mosquitoes make their zooming noise, stimulate their pupils to discuss ‘scientific pre-conceptions’ before moving on to theory. Over the years, projects like Philosophy for Children and Philosophising with Adolescents have been embraced by teachers and the most recent curricula on e.g. Citizenship very much emphasise the need for discussion about and reflection on social and ethical issues. Exploratory talk can also answer the need for 'language developing teaching' tools as part of the Flemish schools' language policy. Last but not least, our pilot study strongly suggests that pupils with weak language skills or those who lack the skills of a Flemish mother tongue speaker can benefit from learning exploratory talk, both linguistically and culturally. For that reason only, further research on this matter is no less than a priority.

Mercer and Dawes (2008) suggest that teachers should facilitate exploratory talk more frequently, especially in the initial stages of a new topic, as this is where pupils are forming and merging concepts. From a pedagogic and organisational point of view this sounds very reasonable. Through exploratory talk teachers can put into practice what Bruner considered essential for effective learning and what Vygotsky believed to be inherent to constructive learning: the critical exploration of foreknowledge and personal experiences (as existing mental structures) in order to clear the path for new insights and knowledge. But if this is to succeed, the ground rules of exploratory talk have to be embedded in the curriculum in a coherent way, making them a *must-teach* for every teacher and insuring a vertical line of development, from kindergarten to higher education. At the same time teachers must be made aware of the fact that exploratory talk is not exclusively tied up to language classes but is a necessary didactic tool which improves learning in every domain. For without a systematic approach, without proper, problem solving assignments and without reflection (feedback and feed forward) on the process of group work/talk itself, learning effects will remain limited, and transfer beyond the context in which exploratory talk was learnt will be very difficult (Wegerif, 1996a).

As Barnes (1976) suggests, high teaching skills are required to put social constructivist teaching into practice. In our country, teacher education focuses very much on all aspects of constructivist learning, including metacognition, but classroom practice tells a more traditional story. From the previous paragraph we know there are several reasons for this, but as far as the Flemish context is concerned, we want to address two problems. The first has to do with teacher educaton. The second is insufficient in-service professionalization.

Student teachers are almost literally - and for a number of topics only once during their education - bombed with valuable learning theories, pedagogic principles, didactic approaches, etc. It is only after a few years of practice that they discover more if not all of their value and relevance. By consequence, it also takes time for them to develop the confidence to walk away from the more traditional paths of their own school career they have - often falsely - been considering as a safe haven for teaching. Therefore, we believe teacher education needs to invest in a more repetitive as well as progressive curriculum which stimulates an continuous synergy between theory, practice and reflection. We propose an inductive and integrated curriculum in which student teachers start - after a basic introduction in general pedagogic theory - with a short and simple apprenticeship, say one week, during which they participate in all kinds of activities in a school, including attending lessons, and during which a basic introduction and observational assignments help them to change their perspective from that of a pupil to that of a teacher. After that, the student teachers' observations can be used as scaffolds for critical reflection and theory-building. Simultaneously general educational and pedagogical theory prepares the students for their next step into practice, which is more demanding: teaching their first lessons, but in a relatively safe context (e.g. by inviting pupils to high school or university, where they are taught in small groups, so that the student teacher can concentrate on didactic approaches and does not have to bother too much about classroom management). After that, reflection and theory use this new experience for further exploration and the deepening of insights. This process is repeated throughout the year(s), while apprenticeships become longer, more complex and more demanding, approaching real life teaching. We believe such a curriculum can embed dialogic teaching and exploratory talk more firmly as it offers the possibility to observe and practice these in various school contexts. Children become better aware of how the exploratory talk could be appropriately applied (as a social mode of thinking) when they learn and practise it in a variety of classroom activities and in various contexts, i.e. school subjects. Doing so, they discover the underlying strategies they need to make the transfer to other contexts (Mercer, 1996; Wegerif, 1996a). We believe the same principle can be applied in the model for teacher education we propose, for what student teachers need is not only good preparation, but also multiple examples of good practice which shows how pupils become enthusiastic and effective at grasping 'educated' ways of using language for sharing and constructing knowledge (Mercer, 1995). Simultaneously, this may feed the initial commitment of students to implement dialogic teaching and exploratory talk in their later practice.

Apart from that, teacher education should not be the end of a student's professional development as a teacher but the beginning. It is very regrettable that teacher education is not prolonged after graduation. It just stops. But even more disturbing is the fact that further training and in-service training are close to problematic in Flemish education. The 2013 TALIS report shows that the average Flemish teacher spends only three days a year on professional development, which is very little compared to other OECD countries (Van Hoof & Van Petegem, 2013). Therefore, we believe it would be a good idea to fully integrate further training systematically in, say, the first three years of any teacher's professional career. At the start of their career teachers need not 'drown' in a full timetable, they should be given space and time to grow into their profession and receive further training, as is common in business. Not only would Flemish teachers 'get the feeling' for professionalisation (and give our country higher scores in the next TALIS report), less young teachers would drop out, as is currently the case. Teacher educators can also benefit, as they would become less separated from daily practice which could, in turn benefit evidence based research (on dialogic teaching, exploratory talk, etc) in schools. Of course, this model also implies that teacher educators should teach as they preach, and apply dialogic teaching during teacher education as well.

The results of this study are not asking to be stored away on a shelf, but to be used and translated to educational practice. Flemish education now has a first scientific ground to embrace exploratory talk as a learning tool. Study 1 has made clear that though the implementation of exploratory talk in curricula may still be problematic, research interest is growing rapidly across the globe. With this study, we have joined this research community and Study 2 has demonstrated the educational potential in our primary school context. As said before, the ground rules are not absent in the Flemish curriculum, but right now they are living a very fragmented and rather obscure life. As a result, they are very unlikely to be perceived as a coherent set of ground rules the implementation of which requires a specific didactic approach. Claiming that Flemish curricula do not pay any attention to oracy would be exaggerated, but we do believe more - and more systematic - emphasis on interaction and conversational strategies are needed.

Flemish education is changing rapidly and one only has to look at the faces of today's pupils to understand what has changed the most: our schools are evolving from basically white institutions to multicultural learning platforms. It is our firm belief that exploratory talk is not only a means to improve learning but also to upgrade cultural understanding among pupils. This dissertation may stimulate schools to actually implement exploratory talk. Unfortunately, where different cultures are brought together, some people seem to prefer emotion to ratio and explosion to exploration. This calls for an answer which teachers can provide. We believe early education of exploratory talk can help children to develop a rational, exploratory reflex in a world in which mutual understanding and sound reflection have never been needed more.

Bibliography

*Articles reviewed for Study 1 are marked with * (Chapter 3, 3.1 Exploratory talk as a concept) and ** (Chapter 3, 3.2 and 3.3 The measurement and effects of exploratory talk)*

Abimo. (2018). Het nieuwe AVI-systeem. Retrieved from http://www.abimo.net/nieuwe-avi-systeem/.

Alexander, R. (2001). *Culture and pedagogy: International comparisons in primary education*. Oxford: Blackwell.

Alexander, R. (2004). Dialogic teaching and the study of classroom talk. In *International Conference Keynote Address,* 44(3). 103-111.

Alexander, R. (2010). *Dialogic teaching essentials*. Singapore: National Institute of Education.

Ardui, J., De Cock, M., R., F., Hinnekint, K., Reybrouck, M. & Schrooten, E. (2011). I*nspiratiegids voor een kwaliteitsvolle vakdidactiek*. Leuven: Expertisenetwerk School of Education.

Asterhan, C. S. (2013). Epistemic and interpersonal dimensions of peer argumentation. *Affective learning together*, 251-271.

*/**Atwood, S., Turnbull, W. & Carpendale, J. (2010). The Construction of Knowledge in Classroom Talk. *Journal of the Learning Sciences, 19*(3), 358-402.

Ausubel, D. P. (1960). The use of advance organisers in the learning and retention of meaningful verbal material. *Journal of educational psychology, 51*(5), 267.

Bakhtin, M. (2010). *The dialogic imagination: Four essays* (Vol. 1). Austin and London: University of Texas Press.

Barbey, A. & Barsalou, L. (2009). Reasoning and problem solving: models. *Encyclopedia of neuroscience, 8*, 35-43.

Barnes, D. (1976). *From communication to curriculum*. Harmondsworth: Penguin Books.

*Barnes, D. (2010). Why talk is important. *English Teaching-Practice and Critique, 9*(2), 7-10.

Barnes, D., Britton, J. & Rosen, H. (1969). *Language, the Learner, and the School*. London: Penguin Books.

Barnes, D. & Todd, F. (1977). *Communication and learning in small groups*. London: Routledge & Kegan Paul Ltd.

Barnes, D. & Todd, F. (1995a). *Communication and learning Revisited*. Portsmouth, NH: Heinemann.

Barnes, D. & Todd, F. (1995b). *Communication and learning revisited: Making meaning through talk.* Porthsmouth, NH.: Boynton/Cook Pub.

**Bayer, A. S. (1986). A case study: Teachers' learning and language development. *Journal of Research & Development in Education, 19*(4), 19-28.

Bazeley, P. & Jackson, K. (2013). *Qualitative data analysis with NVivo*. Thousand Oaks: Sage Publications Limited.

*/**Beghetto, R. A. & Kaufman, J. C. (2009). Intellectual Estuaries: Connecting Learning and Creativity in Programmes of Advanced Academics. *Journal of Advanced Academics, 20*(2), 296-324.

Bentley, A. F. & Dewey, J. (1949). *Knowing and the Known*. Boston: Beacon Press.

*Berkowitz, M. W. & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly (1982)*, 399-410.

Bliss, J., Askew, M. & Macrae, S. (1996). Effective teaching and learning: Scaffolding revisited. *Oxford Review of Education, 22*(1), 37-61.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain* (Vol. 19). New York: David McKay Co Inc.

Bonset, H. & Braaksma, M. (2008). Het schoolvak Nederlands opnieuw onderzocht: Een inventarisatie van onderzoek van 1997 tot en met 2007. Retrieved from www.taalunieversum.org

Bonset, H. & Hoogeveen, M. (2011). Mondelinge taalvaardigheid in het basisonderwijs. Een inventarisatie van empirisch onderzoek. Retrieved from www.taalunieversum.org

Boonen, W. (2006). *Vlaamse normering van de AVI-toets. Onderzoeksinstrument voor het technisch lezen op tekstniveau, niveaubepaling en kwalitatieve analyse*. Leuven: Garant.

Bossaert, S. (2009). *Wie voert het hoogste woord? Interactie in het eerste leerjaar basisonderwijs vanuit genderperspectief.* Brussel: ASP/VUBPRESS/UPA.

*/**Bowskill, N. (2010). Giving Shareable Form to Collective Thought Using a Shared Thinking Approach. Retrieved from https://eric.ed.gov/?id=ED511570.

*/**Boyd, M. & Kong, Y. (2017). Reasoning Words as Linguistic Features of Exploratory Talk: Classroom Use and What It Can Tell Us. *Discourse Processes, 54*(1), 62-81.

**Brevig, L. (2006). Engaging in Retrospective Reflection. *Reading Teacher, 59*(6), 522-630.

Britton, J. (1970). *Language and learning*: Allen Lane London.

Brouwer, T. (2006). Horizontaal en verticaal: een interactieve taalzaakvakles heeft het allemaal. *Vonk, 35*(3), 35-46.

*/**Brown, A. C. (2016). Classroom community and discourse: How argumentation emerges during a Socratic circle. *Dialogic Pedagogy, 4*.

Brown, R. & Renshaw, P. (2000). Collective argumentation: A sociocultural approach to reframing classroom teaching and learning. In H. Cowie & G. van der Aalsvoort (Eds.), *Social interaction in learning and instruction: The meaning of discourse for the construction of knowledge (p.52–66).* New York: Pergamon.

Bruner, J. (1960). *The process of education:[a searching discussion of school education opening new paths to learning and teaching.* London: Vintage Books.

Bruner, J. (1962). *The art of discovery learning. On knowing: Essays for the left hand.* Harvard: Harvard College.

*/**Bullen, K., Moore, K. & Trollope, J. (2002). The Influence of Pupil-Generated Ground Rules on Collaborative Learning in the Classroom: A Pilot Study. *Evaluation and Research in Education, 16*(4), 202-217.

Caroll, J. (1970). *Taal en denken [Language and thought]*. Rotterdam: Universitaire Pers.

Carter, R. (2002). The grammar of talk: Spoken English, grammar and the classroom. *Language Issues, 14*(2), 2-7.

Cazden, C. B. (2001). The language of teaching and learning. *The language of teaching and learning*, 348-369.

**Cervetti, G. N., DiPardo, A. L. & Staley, S. J. (2014). Entering the Conversation: Exploratory Talk in Middle School Science. *Elementary School Journal, 114*(4), 547-572.

Chapman, M. (1991). The epistemic triangle: Operative and communicative components of cognitive competence. In A preliminary version of this chapter was presented at the Meeting of the Society for Research in Child Development, Kansas City, Apr 1989.

*/**Chick, M. (2015). The Education of Language Teachers: Instruction or Conversation? *ELT Journal, 69*(3), 297-307.

Cincinnato, S. & De Meyer, I. (2014). Hoe vaardig is de Vlaming? De resultaten van het PIAAC-onderzoek samengevat. *Welwijs, 25*(1), 16-20.

Coffey, H. (2016). Cooperative learning. Retrieved from www.learnnc.org/lp/pages/4653#note1

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement, 20*(1), 37-46.

Coultas, V. (2016). Case studies of teachers' understandings of the pedagogy of classroom talk: some critical moments explored. *Literacy, 50*(1), 32-39.

*/**Coultas, V. (2012). Classroom talk: Are we listening to teachers' voices? *English in Education, 46*(2), 175-189.

Crook, C. (1998). Children as computer users: the case of collaborative learning. *Computers & Education, 30*(3), 237-247.

Damhuis, R. & De Blauw, A. (2006). Leerlingen aan de praat krijgen: waarom en hoe? Over het interactiegedrag van onderwijsgevenden. *Vonk, 35*(1), 2-17.

Dawes, L. (2004). Talk and learning in classroom science. *International Journal of Science Education, 26*(6), 677-695.

Dawes, L. (2008). *The essential speaking and listening: Talk for learning at Key Stage 2.* Abington: Routledge.

*/**Dawes, L., Dore, B., Loxley, P. & Nicholls, L. (2010). A talk focus for promoting enjoyment and developing understanding in science. *English Teaching-Practice and Critique, 9*(2), 99-110.

Dawes, L., Fisher, E. & Mercer, N. (1992). The quality of talk at the computer. *Language and Learning*, 22-25.

Dawes, L. & Sams, C. (2004). T*alk box: Speaking and listening activities for learning at Key Stage 1*. London, Routledge.

De Corte, E. (1996). Actief leren binnen krachtige onderwijsleeromgevingen. *Impuls, 26*(4), 145-156.

Decoo, W. (2004). Evaluatie in het taalonderwijs: de 40/60-regel doorgelicht. Volledige versie van een ingekorte bijdrage. *Romaneske, 29*(4), 25-47.

Dewey, J. (1933). *How we think: A restatement of the reflective thinking to the educative process*. Boston, C. Heath.

Dillon, J. T. (1988). The remedial status of student questioning. *Journal of Curriculum studies, 20*(3), 197-210.

Dochy, F., Segers, M. & Buehl, M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of educational research, 69*(2), 145-186.

*/**Dourneen, J. (2013). It ain't what you say, it's the way that you say it: The role of pragmatics in A Level English Language students' learning conversations during collaborative writing using a computer. *English in Education, 47*(1), 33-48.

Dudal, P. (2000). *Leerlingvolgsysteem. Spelling: Toetsen 1–6.* Leuven: Garant.

Dudal, P. (2001). *Leerlingvolgsysteem. Wiskunde. (LVS-VCLB).* Antwerpen: Garant.

Edwards, A. D. & Westgate, D. P. (1994). *Investigating classroom talk* (Vol. 13). London: Falmer Press.

Ellis, R. (2012). *Language teaching research and language pedagogy*. Chichester, Malden, MA: John Wiley & Sons.

*/**Enghag, M., Gustafsson, P. & Jonsson, G. (2007). From everyday life experiences to physics understanding occurring in small group work with context rich problems during introductory physics work at university. *Research in Science Education, 37*(4), 449-467.

*/**Enghag, M., Gustafsson, P. & Jonsson, G. (2009). Talking Physics during Small-Group Work with Context-Rich Problems - Analysed from an Ownership Perspective. *International Journal of Science and Mathematics Education, 7*(3), 455-472.

English, E., Hargreaves, L. & Hislam, J. (2002). Pedagogical dilemmas in the National Literacy Strategy: primary teachers' perceptions, reflections and classroom behaviour. *Cambridge Journal of Education, 32*(1), 9-26.

*/**Fernandez, M., Wegerif, R., Mercer, N. & Rojas-Drummond, S. (2001). Re-Conceptualizing "Scaffolding" and the Zone of Proximal Development in the Context of Symmetrical Collaborative Learning. *Journal of Classroom Interaction, 50*(1), 54-72.

Figg, C., Wenrick, M., Youker, C., Heilman, J. & Schneider, C. (2009). Implications and benefits of a long-term peer debriefing experience on teacher researchers. *Brock Education Journal, 19*(1).

**Figueiredo, M. P., Figueiredo, A. & Rego, B. (2015). Peer Interaction in Mixed Aged Groups: A Study in the Computer Area of an Early Childhood Education Center in Portugal. *Iceri2015: 8th International Conference of Education, Research and Innovation*, 2548-2554.

Fisher, A. T. (2011). Creating an articulate classroom: examining pre-service teachers' experiences of talk. *Language and Education, 25*(1), 33-47.

*Fisher, E. (1993). Distinctive Features of Pupil-Pupil Classroom Talk and Their Relationship to Learning: How Discursive Exploration Might Be Encouraged. *Language and Education, 7*(4), 239-257.

Flanders, N. (1970). *Analysing Teaching Behavior*. Reading, Mass.: Addison-Wesly Pub. Co.

Galton, M. & Williamson, J. (1992). *Group work in the Primary Schools*: London, Routledge.

Geudens, V., Rymenans, R. & Daems, F. (1992). Nederlands in de niet-taalvakken. *Eindrapport project FCFO-MI*(90.05). Antwerpen: Universiteit Antwerpen, UFSIA, ICTL.

*Gillies, R. M. (2014). Developments in classroom-based talk. *International Journal of Educational Research, 63*, 63-68.

GO!, O. v. d. V. G. (2013a). Leerplan Nederlands voor het Basisonderwijs. Retrieved from Brussel: http://pro.g-o.be/blog/Documents/Nederlands.pdf

*/**Golanics, J. D. & Nussbaum, E. M. (2008). Enhancing online collaborative argumentation through question elaboration and goal instructions. *Journal of Computer Assisted Learning, 24*(3), 167-180.

**Gordon, J. (2010). Talking about poems, elaborating Barnes. *English Teaching, 9*(2), 47.

Green, S. K. & Gredler, M. E. (2002). A review and analysis of constructivism for school-based practice. *School Psychology Review, 31*(1), 53.

Habermas, J. (1985). *The theory of communicative action* (Vol. 2). Boston: Beacon press.

*/**Haglund, J. & Jeppsson, F. (2012). Using self-generated analogies in teaching of thermodynamics. *Journal of Research in Science Teaching, 49*(7), 898-921.

Halliday, M. A. (1979). Modes of meaning and modes of expression: types of grammatical structure, and their determination by different semantic functions. *Function and context in linguistic analysis*, 57-79.

Harris, M. (1995). Talking in the Middle - Why Writers Need Writing Tutors. *College English, 57*(1), 27-42.

*/**Harris, R. & Ratcliffe, M. (2005). Socio-Scientific Issues and the Quality of Exploratory Talk-What can be Learned from Schools Involved in a "Collapsed Day" Project? *Curriculum Journal, 16*(4), 439-453.

Hart, B. & Risley, T. R. (1995). Meaningful Differences in The Everyday Experience of Young American Children. New York: Brookes.

Hattie, J. (2009). Visible learning: A synthesis of meta-analyses in education. London: Routledge.

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11*(1), 49-76.

Heller, P. & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics, 60*(7), 637-644.

*/**Herrlitz-Biro, L., Elbers, E. & de Haan, M. (2013). Key Words and the Analysis of Exploratory Talk. *European Journal of Psychology of Education, 28*(4), 1397-1415.

*/**Hewitt, E. (2014). Emerging Thoughts on an Approach to Engaging Pupils in Effective Group Talk in Science. *Education 3-13, 42*(4), 402-418.

Hicks, D. (1995-1996). Discourse, Learning, and Teaching. *Review of Research in Education Vol. 21*, 49-95.

Hiebert, J. & Stigler, J. W. (1999). *The teaching gap*. New York: The Free Press.

Hoek, D. (2007). Ontwikkeling van een leeromgeving voor samenwerkend leren. *Pedagogische Studiën, 84*(5), 407.

Hoetker, J. (1968). Teacher questioning behavior in nine Junior High School English Classes. *Research in the Teaching of English, 2, 99-108*.

Holmes, J. (1992). Women's talk in public contexts. *Discourse & Society, 3*(2), 131-150.

Hoogeveen, M. & Bonset, H. (1998). *Het schoolvak Nederlands onderzocht. Een inventaris*. Kessel-Lo: Garant.

Hoogeveen, P. & Winkels, J. (2008). *Het didactische werkvormenboek*. Assen: Uitgeverij Van Gorcum.

Howard-Jones, P. (2010). Introducing neuroeducational research. Neuroscience, education and the brain from contexts to practice. Londen/New York: Routledge.

Huddleston, T. & Rowe, D. (2003). Citizenship and the role of language. *Learning to teach citizenship in the secondary school (London: Routledge Falmer)*, 111-130.

**Hunter, R. (2008). Do they know what to ask and why? Teachers shifting student questioning from explaining to justifying and generalising reasoning. *Proceedings of PME 32 and PME-NA 30, 3, 201-208.*

**Hyun, E., & Davis, G. (2005). Kindergartners' conversations in a computer-based technology classroom. *Communication education, 54*(2), 118-135.

IJzendoorn, T. v. (1998). Begrijpend lezen door rolwisselend leren: Lisse: Swets en Zeitlinger.

**Jacobs, S. E., & Karliner, A. B. (1977). Helping writers to think: The effect of speech roles in individual conferences on the quality of thought in student writing. *College English, 38*(5), 488-505.

Janssen, J., Kirschner, F., Erkens, G., Kirschner, P. A. & Paas, F. (2010). Making the black box of collaborative learning transparent: Combining process-oriented and cognitive load approaches. *Educational psychology review, 22*(2), 139-154.

Jay, T., Taylor, R., Moore, N., Burnett, C., Merchant, G., Thomas, P., et al. (2017). *Dialogic Teaching: evaluation report and executive summary.* London: Education Endowment Foundation with Sheffield Hallam University. https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation-Reports/Dialogic-Teaching-Evaluation-Report. pdf.

Jonassen, D. H. & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research and Development, 58*(4), 439-457.

Jones, K. (2004). Mission Drift in Qualitative Research, or Moving Toward a Systematic Review of Qualitative Studies, Moving Back to a More Systematic Narrative Review. *The Qualitative Report, 9*(1), 95-112.

Kagan, S. (2014). Coöperatieve leerstrategieën: research, principes en de praktische uitwerking: beknopte uitgave. Rotterdam: Bazalt.

Kassoti, O. & Kliapis, P. (2009). Exploratory Talk for Problem Solving. Pme 33: Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education, Vol 3, 289-296.

Kaufer, D. (2011). What can Neuroscience Research Teach Us about Teaching? Retrieved from http://gsi.berkeley.edu/programmes-services/hsl-project/hsl-speakers/kaufer/

**Kerawalla, L. (2015). Talk Factory Generic: Empowering secondary school pupils to construct and explore dialogic space during pupil-led whole-class discussions. *International Journal of Educational Research, 70*, 57-67.

**Kerawalla, L., Petrou, M. & Scanlon, E. (2013). Talk Factory: Supporting "Exploratory Talk" around an Interactive Whiteboard in Primary School Science Plenaries. *Technology, Pedagogy and Education, 22*(1), 89-102.

**Kerr, J.-A. (1998). Listening in on the Language of Collaboration in a Writing Workshop. *Middle School Journal, 30*(1), 3-8.

Kirkevold, M. (1997). Integrative nursing research - an important strategy to further the development of nursing science and Practice. *Journal of Advanced Nursing, 25*, 977-984.

*/**Knight, S. & Mercer, N. (2015). The role of exploratory talk in classroom search engine tasks. *Technology, Pedagogy and Education, 24*(3), 303-319.

Köhler, W. (1970). Gestalt psychology: An introduction to new concepts in modern psychology. New York: WW Norton & Company.

*/**Kruger, A. C. (1993). Peer collaboration: Conflict, cooperation, or both? *Social Development, 2*(3), 165-182.

**Kucirkova, N., Messer, D., Sheehy, K. & Panadero, C. F. (2014). Children's engagement with educational iPad apps: Insights from a Spanish classroom. *Computers & Education, 71*, 175-184.

Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. M. K., Hammerstrøm, K., & Sathe, N. (2016). *Searching for Studies: A Guide to Information Retrieval for Campbell Systematic Reviews.* Oslo: The Campbell Collaboration.

*/**Kumpulainen, K. (1996). The nature of peer interaction in the social context created by the use of word processors. *Learning and Instruction, 6*(3), 243-261.

**Kumpulainen, K. & Wray, D. (1999). Analysing Interactions during Collaborative Writing with the Computer: An Innovative Methodology. Retrieved from: http://www.warwick.ac.uk/staff/D.J.Wray/Articles/facct.html

Lammers, H. (1993). De impasse in het onderwijs in de mondelinge vaardigheden. *Spiegel, 3*, 9-35.

Lefstein, A. (2008). Changing classroom practice through the English National Literacy Strategy: A micro-interactional perspective. *American Educational Research Journal, 45*(3), 701-737.

Leidse Werkgroep Moedertaaldidactiek (1982). *Moedertaalonderwijs in ontwikkeling*. Muiderberg: Dick Coutinho.

Lemke, J. (1990). *Talking science: Language, learning and values*. Norwood: Ablex Publishing Company.

Light, P. (1991). Peers, problem solving, and computers. *Golem, 3*(1), 2-6.

Lipman, M. (1981). Philosophy for children. *Developing minds: Programmes for teaching thinking, 2*, 35-38.

Littleton, K. & Mercer, M. (2010). The Significance of Educational dialogues between primary school children. Teoksessa K. Littleton & C. Howe (Eds.) *Educational dialogues: understanding and promoting productive interaction.* (pp. 48-63). London: Routledge.

Littleton, K. & Mercer, N. (2013). *Interthinking: Putting talk to work.* London: Routledge.

Littleton, K., Mercer, N., Dawes, L., Wegerif, R., Rowe, D. & Sams, C. (2005). Talking and thinking together at Key Stage 1. *Early years, 25*(2), 167-182.

*Lofgren, R., Schoultz, J., Hultman, G. & Bjorklund, L. (2013). Exploratory Talk in Science Education: Inquiry-Based Learning and Communicative Approach in Primary School. *Journal of Baltic Science Education, 12*(4), 482-496.

*/**Luby, A. (2014). First Footing Inter-Faith Dialogue. *Educational Action Research, 22*(1), 57-71.

*Lyle, S. (2008). Dialogic teaching: Discussing theoretical contexts and reviewing evidence from classroom practice. *Language and Education, 22*(3), 222-240.

**Makgato, M. (2015). Barriers associated with the use of English in the teaching of technology in grade 9 at some schools in Eastern Cape Province. *Africa Education Review, 12*(2), 180-192.

**Maloch, B. (2002). Scaffolding student talk: One teacher's role in literature discussion groups. *Reading Research Quarterly, 37*(1), 94-112.

*/**Mannion, J. & Mercer, N. (2016). Learning to learn: improving attainment, closing the gap at Key Stage 3. *Curriculum Journal, 27*(2), 246-271.

Mayher, J. S. (1990). *Uncommon sense: Theoretical practice in language education*: Portsmouth: Heinemann Educational Publishers.

McConaghy, J. L. (2014). The Centrality of Exploratory Talk in Dialogic Teaching and Learning. Alberta: University of Alberta.

Mehan, H. (1979). *Learning lessons.* Cambridge, MA: Harvard University Press.

Mercer, N. (1995). The guided construction of knowledge: Talk amongst teachers and learners. Bristol: Multilingual matters.

**Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction, 6*(4), 359-377.

Mercer, N. (2000). *Words and minds: How we use language to think together*: Psychology Press. London and New York: Routledge.

**Mercer, N. (2004). Sociocultural discourse analysis. *Journal of applied linguistics, 1*(2), 137-168.

**Mercer, N. (2010a). The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology, 80*(1), 1-14.

Mercer, N. (2010b, 25 November 2010). *Language and education: improving pupils' skills in communication, learning and reasoning.* Paper presented at the Leerstoel Onderwijsvernieuwing en -samenwerking, Antwerp University.

Mercer, N. (2013). The Social Brain, Language, and Goal-Directed Collective Thinking: A Social Conception of Cognition and Its Implications for Understanding How We Think, Teach, and Learn. *Educational Psychologist, 48*(3), 148-168.

Mercer, N. & Hodgkinson, S. (2008). *Exploring talk in school: Inspired by the work of Douglas Barnes.* Thousand Oaks: Sage Publications Ltd.

Mercer, N. & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction, 1*(1), 12-21.

Mercer, N. & Littleton, K. (2007). Dialogue and the development of children's thinking: A sociocultural approach. London: Routledge.

**Mercer, N. & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education, 20*(6), 507-528.

Mercer, N., Hodgkinson, S., & Alexander, R. J. (2008). Culture, dialogue and learning: notes on an emerging pedagogy. In *Exploring Talk in School: Inspired by the Work of Douglas Barnes* (pp. 91-114). London: Sage.

**Mercer, N., Littleton, K. & Wegerif, R. (2004). Methods for studying the processes of interaction and collaborative activity in computer-based educational activities. *Technology, Pedagogy and Education, 13*(2), 195-212.

**Mercer, N., Warwick, P., Kershner, R., & Staarman, J. K. (2010). Can the interactive whiteboard help to provide ‘dialogic space’for children's collaborative activity?. *Language and education, 24*(5), 367-384.

*/**Mercer, N., Wegerif, R. & Dawes, L. (1999). Children's Talk and the Development of Reasoning in the Classroom. *British Educational Research Journal, 25*(1), 95-111.

*Michaels, S., O’Connor, C. & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in philosophy and education, 27*(4), 283-297.

Miller, J. E., Trimbur, J. & Wilkes, J. M. (1994). Group dynamics: Understanding group success and failure in collaborative learning. *New directions for teaching and learning, 1994*(59), 33-44.

Mortimer, E. & Scott, P. (2003). *Meaning Making in Science Classrooms*. Milton Keynes: Open University Press.

*/**Murcia, K. & Sheffield, R. (2010). Talking about science in interactive whiteboard classrooms. *Australasian Journal of Educational Technology, 26*(4), 417-431.

**Murphy, C. (2015). Changing Teachers' Practices through Exploratory Talk in Mathematics: A Discursive Pedagogical Perspective. *Australian Journal of Teacher Education, 40*(5).

Myhill, D. & Dunkin, F. (2005). Questioning learning. *Language and Education, 19*(5), 415-427.

*/**Nikolaidou, G. N. (2012). ComPLuS model: A new insight in pupils' collaborative talk, actions and balance during a computer-mediated music task. *Computers & Education, 58*(2), 740-765.

**Nussbaum, E. M. (2005). The Effect of Goal Instructions and Need for Cognition on Interactive Argumentation. *Contemporary Educational Psychology, 30*(3), 286-313.

*Nussbaum, E. M. (2004). Goal instructions in computer-supported collaborative argumentation: Can you find a reason? *Icls2004: International Conference of the Learning Sciences, Proceedings*, 380-387.

*Nussbaum, E. M. (2015). The Effect of Goal Instructions and Need for Cognition on Interactive Argumentation. *Contemporary Educational Psychology, 30*(3), 286-313.

**O'Donnell-Allen, C. (2001). Teaching with a Questioning Mind: The Development of a Teache Research Group into a Discourse Community. *Research in the Teaching of English*, 161-211.

Ochs, E. (1979). Transcription as theory. *Developmental pragmatics, 10*(1), 43-72.

OECD, 2016. *PISA 2015 Results (Volume I)*. Paris: OECD Publishing.

Paelman, F. (2004). *CLIM-wijzer voor het secundair onderwijs*. Wommelgem. De Boeck.

Palincsar, A. S. & Brown, A. L. (1986). Interactive teaching to promote independent learning from text. *The Reading Teacher, 39*(8), 771-777.

Panitz, T. (1997). Why more teachers do not use collaborative learning techniques. Retrieved from http://home.capecod.net/~tpanitz/tedsarticles/whyfewclusers.

**Pantaleo, S. (2011). Middle years students' collaborative talk about'The Red Tree':'A book that really works your mind'. *Australian Journal of Language and Literacy, 34*(3), 260.

Paus, H., Rymenans, R. & Van Gorp, K. (2006). *Dertien doelen in een dozijn. Een referentiekader voor taal*. Den Haag: Nederlandse Taalunie.

Perry, W. G. (1999). Forms of Intellectual and Ethical Development in the College Years: A Scheme. Jossey-Bass Higher and Adult Education Series. ERIC.

Philips, S. U. (2001). Participant structures and communicative competence: Warm Springs children in community and classroom. New York: Longman.

Phillips, T. (1992). Why? The neglected question in planning small group activity. *Thinking voices: The work of the national oracy project. London: Hodder and Stoughton*, 148-155.

Piaget, J. (1970). Science of education and the psychology of the child. Trans. D. Coltman. Washington: American Psychological Association.

Piaget, J. (1995). *Sociological studies.* (Original work published 1977). London: Psychology Press.

Piaget, J. & Elkind, D. (1968). *Six psychological studies* (Vol. 462). London: Vintage Books.

**Pierce, K. M. (1995). A Plan for Learning: Creating a Place for Exploratory Talk. *Primary Voices K-6, 3*(1), 16-29.

*/**Polo, C., Lund, K., Plantin, C. & Niccolai, G. (2015). *Analysing Exploratory Talk as a Socio-Cognitive Practice: Identity, Group Argumentation, and Class Debate Quality.* HAL Archives Ouvertes. Retrieved from: https://hal.archives-ouvertes.fr/hal-01208319/document.

Pressley, M. & Allington, R. L. (2014). *Reading instruction that works: The case for balanced teaching*. New York: Guilford Publications.

Rabel, S. & Wooldridge, I. (2013). Exploratory talk in mathematics: what are the benefits? *Education 3-13, 41*(1), 15-22.

*/**Rajala, A., Hilppo, J. & Lipponen, L. (2012). The Emergence of Inclusive Exploratory Talk in Primary Students' Peer Interaction. *International Journal of Educational Research, 53*, 55-67.

Raven, J., raven, J.C. & Court, J.H. (1998). Manual for Raven’s progressive matrices and vocabulary scales. *Oxford: Oxford Psychologists*.

**Reninger, K. B. & Rehark, L. (2009). Discussions in a Fourth-Grade Classroom: Using Exploratory Talk to Promote Children's Dialogic Identities. *Language Arts, 86*(4), 268-279.

**Renshaw, P. & Brown, R. A. (2007). Formats of classroom talk for integrating everyday and scientific discourse: Replacement, interweaving, contextual privileging and pastiche. *Language and Education, 21*(6), 531-549.

*/**Riley, N. R. (2006). Methods for Evaluating Critical Learning Using Online Discussion Forums. *Technology, Pedagogy and Education, 15*(1), 63-78.

*/**Robins, G. (2011). The effect of exploratory talk on the development of sentence structure in able writers. *Literacy, 45*(2), 78-83.

Rockwood III, H. (1995). *Cooperative and collaborative learning.* Paper presented at the The national teaching & learning forum.

*/**Rojas-Drummond, S., Mazon, N., Fernandez, M. & Wegerif, R. (2006). Explicit Reasoning, Creativity and Co-Construction in Primary School Children's Collaborative Activities. *Thinking Skills and Creativity, 1*(2), 84-94.

*/**Rojas-Drummond, S. & Mercer, N. (2003). Scaffolding the development of effective collaboration and learning. *International Journal of Educational Research, 39*(1), 99-111.

*/**Rojas-Drummond, S., Mercer, N. & Dabrowski, E. (2001). Collaboration, scaffolding and the promotion of problem solving strategies in Mexican pre-schoolers. *European Journal of Psychology of Education, 16*(2), 179-196.

*/**Rojas-Drummond, S., Perez, V., Velez, M., Gomez, L. & Mendoza, A. (2003). Talking for reasoning among Mexican primary school children. *Learning and Instruction, 13*(6), 653-670.

*/**Rojas-Drummond, S., Torreblanca, O., Pedraza, H., Velez, M. & Guzman, K. (2013). 'Dialogic scaffolding': Enhancing learning and understanding in collaborative contexts. *Learning Culture and Social Interaction, 2*(1), 11-21.

*/**Rojas-Drummond, S. & Zapata, M. P. (2004). Exploratory Talk, Argumentation and Reasoning in Mexican Primary School Children. *Language and Education, 18*(6), 539-557.

*/**Rutter, T., Edwards, R. & Dean, P. (2016). Who's that talking in my class?: What does research say about pupil to pupil exploratory talk that leads to learning? *Teacher Education Advancement Network Journal, 8*(1), 22-32.

Sandra, D., Frisson, S. & Daems, F. (1999). Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in Dutch. *Brain and language, 68*(1-2), 277-283.

Schittekatte, M. (2000). Het hertesten van intelligentie bij kinderen: een literatuurstudie. Retrieved from: http://www.sig-net.be/uploads/artikels signaal/significant hertesting schittekatte 2000 nr0.pdf.

**Schwartz, M., & Gorbatt, N. (2016). ‘Why do we know Hebrew and they do not know Arabic?’ Children’s meta-linguistic talk in bilingual preschool. *International Journal of Bilingual Education and Bilingualism, 19*(6), 668-688.

*/**Schmitz, M. J. & Winskel, H. (2008). Towards effective partnerships in a collaborative problem-solving task. *British Journal of Educational Psychology, 78*, 581-596.

Scott, P. H., Mortimer, E. F. & Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education, 90*(4), 605-631.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language* (Vol. 626). Cambridge: Cambridge University Press.

**Setati, M., Adler, J., Reed, Y. & Bapoo, A. (2002). Incomplete Journeys: Code-Switching and Other Language Practices in Mathematics, Science, and English Language Classrooms in South Africa. *Language and Education, 16*(2), 128-149.

Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). Construct validity and external validity. *Experimental and quasi-experimental designs for generalised causal inference*, 64-102.

Sheeran, Y. & Barnes, D. R. (1991). *School writing: Discovering the ground rules.* Michigan: Open University Press.

Sinclair, J. & Coulthard, M. (1977). Towards an Analysis of Discourse: The English Used by Teachers and Pupils. *TESOL Quarterly, 11(2)*, 203-206.

**Skidmore, D. & Murakami, K. (2010). How Prosody Marks Shifts in Footing in Classroom Discourse. *International Journal of Educational Research, 49*(2-3), 69-77.

Skinner, B. F. (1972). *Beyond freedom and dignity*. London: Springer.

Slavin, R. E. (1996). Cooperative learning in middle and secondary schools. *The Clearing House, 69*(4), 200-204.

*/**Soter, A. O., Wilkinson, I. A., Murphy, P. K., Rudge, L., Reninger, K. & Edwards, M. (2008). What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research, 47*(6), 372-391.

*/**Sutherland, J. (2006). Promoting Group Talk and Higher-Order Thinking in Pupils by "Coaching" Secondary English Trainee Teachers. *Literacy, 40*(2), 106-114.

*/**Sutherland, J. (2013). Going "Meta": Using a Metadiscoursal Approach to Develop Secondary Students' Dialogic Talk in Small Groups. *Research Papers in Education, 30*(1), 44-69.

**Sweigart, W. (1991). Classroom talk, knowledge development, and writing. *Research in the Teaching of English*, 469-496.

T'Sas, J. (2013). *Sprekend leren. Wanneer groepswerk echt rendeert in het lager onderwijs.* Brugge: Die Keure.

Thomson, A. J. & Martinet, A. V. (1986). *A practical English grammar.* Oxford: Oxford University Press.

**Tin, T. B. (2003). Does talking with peers help learning? The role of expertise and talk in convergent group discussion tasks. *Journal of English for Academic Purposes, 2*(1), 53-66.

Tokuhama-Espinosa, T. (2010). Mind, brain, and education science: A comprehensive guide to the new brain-based teaching. New York: dWW Norton & Company.

Topolinski, S. & Reber, R. (2010). Gaining insight into the “Aha” experience. *Current Directions in Psychological Science, 19*(6), 402-405.

Topping, K., Campbell, J., Douglas, W. & Smith, A. (2003). Cross-age peer tutoring in mathematics with seven-and 11-year-olds: Influence on mathematical vocabulary, strategic dialogue and self-concept. *Educational Research, 45*(3), 287-308.

**Topping, K. J. & Trickey, S. (2014). The role role of dialog in philosophy for children. *International Journal of Educational Research, 63*, 69-78.

Valcke, M. (2010). *Onderwijskunde als ontwerpwetenschap*. Gent: Academia Press.

van Boxtel, C., van der Linden, J. & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction, 10*(4), 311-330.

Van Camp, T., Vloeberghs, L. & Tijtgat, P. (2015). *Krachtig leren, cognitief neurowetenschappelijk benaderd*. Leuven: Acco.

van Daal, T., Lesterhuis, M., Coertjens, L., Donche, V. & De Maeyer, S. (2016). Validity of comparative judgement to assess academic writing: examining implications of its holistic character and building on a shared consensus. *Assessment in Education: Principles, Policy & Practice*, 1-16.

Van den Branden, K. (2012). *Handboek taalbeleid basisonderwijs*. Leuven: Acco.

Van den Branden, K. (2015). *Onderwijs voor de 21ste eeuw. Een boek voor leerkrachten en ouders*. Leuven: Acco.

Van den Branden, K. & Van Gorp, K. (2000). How to evaluate CLIM in terms of intercultural education. *Intercultural education, 11*(sup1), 42-51.

van der Aalsvoort, M. & van der Leeuw, B. (1982). *Leerlingen, taal en school; de rol van taal in elke onderwijsleersituatie*. Enschede: SLO..

van der Aalsvoort, M. & van der Leeuw, B. (1992). *Taal, school en kennis*. Enschede: VALO-Moedertaal en Nederlandse Taalunie.

van der Veen, C., Dobber, M., & van Oers, H. J. M. (2017). *Dialogic classroom talk and the dialogical self.* Paper presented at International Society of Cultural and Activity Research (ISCAR), Quebec, Canada.

van Gelderen, A. (1994). *Taalvaardigheidseisen in het zaakvakonderwijs voor eentalige en meertalige klassen; moeilijkheden in de instructietaal en leerdoelen voor het taalonderwijs op de basisschool*. Amsterdam: SCO.

Van Gorp, K. (2010). *# Meester, wat is genen?# De effectiviteit van kennisconstructie en schooltaalverwerving bij NT2-leerlingen aan het einde van het basisonderwijs.* Leuven: Centre for Language and Education.

Van Hoof, J. & Van Petegem, P. (2013). *Talis 2013: Teaching and learning international survey: Professionele ontwikkeling en samenwerking van leraren en schoolleiders in Vlaanderen.* Antwerpen: Universiteit Antwerpen.

Van Lier, L. (1996). *Interaction in the language curriculum: awareness, autonomy and authenticity*. London: Longman.

Van Petegem, P., Imbrecht, I. & Brandt, W. (2005). *Leer- en onderwijspsychologische theorieën*. Gent: Academia Press.

Vandenbroucke, F. (2007). De lat hoog voor talen in iedere school. Goed voor de sterken, sterk voor de zwakken (Setting the bar high for languages in every school. Good for the strong, strong for the weak). Policy Paper Flemish Minister of Education and Training.

Vermunt, J. D. H. M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs: Naar procesgerichte instructie in zelfstandig denken*. Lisse: Swerts and Zeitlinger.

Vlaams Parlement. (1998, 1998-02-16). Over de VEV-nota "Het Onderwijs in Vlaanderen: onderwijs in de jaren negentig en het belang in de toekomst. De onderwijsvisie van het Vlaams Economisch Verbond". (942-1). Brussel: Vlaams Parlement.

Vlaams Verbond van het Katholiek Onderwijs (2013). Leerplannen. Brussel. Retrieved from: http://www.katholiekonderwijs.vlaanderen/

Vlaams Ministerie van Onderwijs en Vorming. (2015). *Onderwijsspiegel 2015. Jaarlijks rapport van de onderwijsinspectie*. Brussel.

Vlaamse overheid. (2004). Peiling informatieverwerving en -verwerking in de eerste graad secundair onderwijs (A-stroom). Brussel. Retrieved from: http://eindtermen.vlaanderen.be/peilingen/index.htm

Vlaamse overheid. (2008). Peiling Nederlands in de derde graad algemeen, technisch en kunstsecundair onderwijs. Brussel. Retrieved from: http://eindtermen.vlaanderen.be/peilingen/index.htm

Vlaamse overheid. (2010). Peiling Nederlands in de derde graad algemeen, technisch en kunstsecundair onderwijs. Brussel. Retrieved from: http://eindtermen.vlaanderen.be/peilingen/index.htm

Vlaamse overheid. (2011). Onderwijsspiegel schooljaar 2009-2010. Verslag over de toestand van het onderwijs. Brussel. Retrieved from: http://eindtermen.vlaanderen.be/peilingen/index.htm

Vlaamse overheid. (2014). Peiling lezen en luisteren in het basisonderwijs. Brussel. Retrieved from: http://eindtermen.vlaanderen.be/peilingen/index.htm

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process.* Cambridge, MA: Harvard University Press.

**Walsh, A., & Yeoman, E. (1999). Making sense of the French in French immersion: Concept development in early FI. *Canadian modern language review, 55*(3), 339-354.

Watson, J. B. (1920). Is thinking merely action of language mechanisms? *British Journal of Psychology, 11*(1), 87-104.

Watson, K. (1985). Some Aspects of Classroom Discourse in English. *Highway One, 8*(1-2), 114-123.

Watson, K. & Young, B. (1986). Discourse for learning in the classroom. *Language Arts, 63*(2), 126-133.

*Webb, L. (2015). The Use of Cartoons as a Tool to Support Teacher Ownership of Mathematics Curriculum Change. *African Journal of Research in Mathematics Science and Technology Education, 19*(1), 57-68.

Webb, N. M., Nemer, K. M. & Ing, M. (2006). Small-group reflections: Parallels between teacher discourse and student behavior in peer-directed groups. *The Journal of the Learning Sciences, 15*(1), 63-119.

*/**Webb, P. & Treagust, D. F. (2006). Using Exploratory Talk to Enhance Problem-Solving and Reasoning Skills in Grade-7 Science Classrooms. *Research in Science Education, 36*(4), 381-401.

*/**Webb, P., Whitlow, J. & Venter, D. (2016). From Exploratory Talk to Abstract Reasoning: a Case for Far Transfer? *Educational psychology review*, 1-17.

*/**Wegerif, R. (1996a). Collaborative learning and directive software. *Journal of Computer Assisted Learning, 12*(1), 22-32.

*/**Wegerif, R. (1996b). Using computers to help coach exploratory talk across the curriculum. *Computers & Education, 26*(1-3), 51-60.

*Wegerif, R. (2005). Reason and Creativity in Classroom Dialogues. *Language and Education, 19*(3), 223-237.

*Wegerif, R. (2008). Dialogic or Dialectic? The Significance of Ontological Assumptions in Research on Educational Dialogue. *British Educational Research Journal, 34*(3), 347-361.

Wegerif, R. (2013). *Dialogic: Education for the Internet age.* London: Routledge.

Wegerif, R., Fujita, T., Doney, J., Linares, J. P., Richards, A. & Van Rhyn, C. (2017). Developing and trialing a measure of group thinking. *Learning and Instruction, 48*, 40-50.

*/**Wegerif, R., Linares, J. P. & Rojas-Drummond, S. (2005). Thinking Together in the UK and Mexico: Transfer of an Educational Innovation. *Journal of Classroom Interaction, 40*(1), 40-48.

Wegerif, R., Littleton, K. & Jones, A. (2003). Stand-alone computers supporting learning dialogues in primary classrooms. *International Journal of Educational Research, 39*(8), 851-860.

*/**Wegerif, R. & Mercer, N. (1997a). A dialogical framework for researching peer talk. *Language and Education Library, 12*, 49-64.

**Wegerif, R. & Mercer, N. (1997b). Using computer-based text analysis to integrate qualitative and quantitative methods in research on collaborative learning. *Language and Education, 11*(4), 271-286.

**Wegerif, R., Mercer, N. & Dawes, L. (1998). Software design to support discussion in the primary curriculum. *Journal of Computer Assisted Learning, 14*(3), 199-211.

*/**Wegerif, R., Mercer, N. & Dawes, L. (1999). From Social Interaction to Individual Reasoning: An Empirical Investigation of a Possible Socio-Cultural Model of Cognitive Development. *Learning and Instruction, 9*(6), 493-516.

Wegerif, R., Mercer, N. & Rojas-Drummond, S. (1999). Language for the social construction of knowledge: Comparing classroom talk in Mexican preschools. *Language and Education, 13*(2), 133-150.

*Westgate, D., & Hughes, M. (2015). Speaking and listening across the primary curriculum: an entry to improved learning and a focus for CPD. *Education 3-13, 43(*5), 565-578.

**Wheeldon, I. (2006). Peer Talk. Mathematics Teaching Incorporating Micromath, 199, 39-41.

Willis, J. (2007). Review of research: Brain-based teaching strategies for improving students' memory, learning, and test-taking success. *Childhood Education, 83*(5), 310-315.

Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry, 17*(2), 89-100.

**Yaguchi, M., Iyeiri, Y. & Baba, Y. (2010). Speech style and gender distinctions in the use of very and real/really: An analysis of the Corpus of Spoken Professional American English. *Journal of Pragmatics, 42*(3), 585-597.

Zhang, Z. & Lu, J. (2014). Assessing Argumentative Representation with Bayesian Network Models in Debatable Social Issues. *International Education Studies, 7*(11), 120-135.

Appendices

1. Pedagogic materials

Pedagogic materials include the basic lessons for exploratory talk, an example of the ground rules poster, the talking prompts and the assessment forms used for feedback and feed forward reflection.

1.1 Basic lessons (Dutch)

Basisles 1

Doelen • De leerlingen leren wat praten zoal inhoudt. Succescriterium • Ik kan aan iemand anders vertellen hoe we zoal praten	Verdieping activiteit Focus op • Vragen stellen, het verband tussen spreken en luisteren

Materialen

- ✓ Babbelkaartjes (‘Sprekend leren’ p. 61 of website) per kind, geknipt en gelamineerd
- ✓ Kaartjes met praattekeningen (‘Sprekend leren’ p. 68 of website)
- ✓ Kopie per groep uit Tan, Shaun (2008) ‘ *De aankomst*’ Sint Niklaas, Querido
- ✓ Kopie uit Joke Van Leeuwen (1989) ‘*We zijn allang begonnen maar nu begint het echt*.’ Sint-Niklaas, Querido
- ✓ Tekstballonnen (groot genoeg om in te schrijven)
- ✓ Post-its in verschillende kleuren
- ✓ Grote gekleurde bladen A2 voor aan het bord (of schema op het bord)
- ✓ Gekleurd schema per groep A3
- ✓ Leue, D. (2009). *Appartemensen*. Tielt, Lannoo. Track 8
- ✓ Kopie A4 uit Leue, D. (2009). *Appartemensen*. Tielt, Lannoo. P. 48-49 en kopieën p. 50, 53, 55, 57.
- ✓ Schrijfstroken personages Bonnie – Polluc uit Leue, D. (2009). *Appartemensen*. Tielt, Lannoo

Tijdens de eerste twee basislessen wordt er klassikaal en in groepjes gewerkt met een kleurenschema. Dit schema wordt bewaard en kan in latere basislessen of groepswerk terug opgehangen worden als ondersteuning of uitbreiding bij exploratieve gesprekken.

Ik	Ik en jij
Waarom praten?	Wanneer is praten Lastig?

Klassikaal werk

Jijzelf	Ik en jij
Waarom praten?	Wanneer is praten Lastig?

1. JIJZELF

Je hangt het blauwe vak van het kleurenschema vooraan in de klas.

Je deelt 1 tekstballon uit per leerling.

Opdracht:

- De leerlingen schrijven individueel op een tekstballon een woord dat ze zelf bedacht hebben toen ze leerden praten.
- De leerkracht verzamelt de uitspraken in het blauwe vak (met foto van de leerlingen?), de leerlingen lichten heel kort toe wat hun woord betekent en hoe het ontstaan is.
- De leerlingen vertellen aan elkaar hoe ze leerden praten en hoe ze denken dat iemand leert praten.
- Eventueel bouwt de leerkrachte extra ervaring in via een spelletje hints of gebarentaal.
- ⇨ Hoe leer je praten?
- ⇨ Waarom is het goed is dat je kunt praten?

Jijzelf	Ik en jij
Waarom praten?	Wanneer is praten Lastig?

2. JIJ en DE ANDER

De leerkracht voert met de leerlingen een klasgesprek. Eventueel laat hij de leerlingen eerst vragen beantwoorden op post-its. Deze antwoorden vormen het uitgangspunt voor een klasgesprek.

Onderstaande vragen komen aan bod. Per vraag gebruikt de leerkracht een post-it in een andere kleur zodat hij ze daarna gemakkelijk kan groeperen. De antwoorden worden klassikaal besproken en opgesomd of verzameld (post-its) in het rode vak van het schema.

- ⇨ Wie kan veel en heel goed praten volgens jou?
- ⇨ Wie is volgens jou eerder stil?
- ⇨ Met wie praat je graag?
- ⇨ Waarover praat je graag met vrienden?
- ⇨ Wat is een babbelkous?

Jijzelf	Ik en jij
Waarom praten?	Wanneer is praten Lastig?

3. HET NUT VAN PRATEN

De leerkracht laat de leerlingen kennismaken met een afbeelding uit ‘De aankomst’ van Shaun Tan (niet opgenomen in dit proefschrift wegens copyright). Hij kan ook verwijzen naar migratie in de actualiteit.

De leerlingen bespreken in groep wat ze denken, voelen, verwachten bij de afbeeldingen. De volgende vragen kunnen hen hierbij helpen:

- ⇨ Wat gebeurt er? Heb je dit zelf al eens meegemaakt?
- ⇨ Hoe los(te) je dat op?
- ⇨ Heb je taal echt nodig?
- ⇨ Wat kun je doen met taal, wat kun je doen door te praten?
- ⇨ Welke dingen kan je beter doen als je er met elkaar over praat?
- ⇨ Wat kan je op school doen door met elkaar te praten?

Tijdens een klassikale nabespreking noteert de leerkracht de antwoorden van de leerlingen in het groene vak op het bord.

Jijzelf	Ik en jij
Waarom praten?	Wanneer is praten Lastig?

4. WANNEER IS PRATEN MOEILIJK?

De leerkracht verwijst terug naar 'De aankomst' of werkt met een illustratie uit 'Wij al lang begonnen maar nu begint het echt' van Joke Van Leeuwen (niet opgenomen in dit proefschrift wegens copyright).

De leerkracht bespreekt klassikaal de volgende vragen en noteert de antwoorden in het blauwe vak:

- ⇨ Wanneer is het moeilijk om met mensen te praten?
- ⇨ Wanneer is het beter dat je niet praat?

Groepswerk (groepjes van 3, met verslaggever)

→ Succescriterium: 'Ik kan aan iemand anders vertellen hoe we zoal praten.'

Opdracht 1

De leerlingen krijgen een lijst met taalhandelingen + post –its. Ze krijgen ook het kleurenschema zoals gebruikt in klassikaal moment 1. Ze bespreken samen de taalhandelingen, schrijven ze op een post-it en proberen ze een plaats te geven in het kleurenschema. Deze taalhandelingen kunnen in verschillende vakken tegelijk geordend worden, bv. argumenteren: hoort bij ik-jij , ik gebruik het op school maar ook met vrienden om toe te lichten waarom ik voor iets kies.

Taalhandelingen:

- iemand overhalen
- roddelen
- argumenteren
- uitleggen
- vragen stellen
- vertellen
- een compliment geven
- iemand afbreken

- je verontschuldigen
- bedanken
- liegen
- waarschuwen
- uitnodigen
- troosten
- iets beweren
- vloeken
- iemand doen lachen

Opdracht 2

Overleg wat de mensen doen, zeggen, wat er verder gaat gebeuren.

Materiaal

- ✓ Situatiekaartjes ('Sprekend leren' p. 68)

Klassikaal moment:

De leerkracht overloopt klassikaal de resultaten van het groepswerk. Hierbij kun hij het grote schema gebruiken om de resultaten te situeren. De volgende vragen komen aan bod:

- ⇨ Wat ging goed?
- ⇨ Wie had goede ideeën?
- ⇨ Wie kon er goed luisteren?
- ⇨ Wat zijn de voordelen van werken in groep?
- ⇨ Waarom laten leraren leerlingen in groepen werken omdat je , precies door met elkaar te praten, ideeën aan elkaar kunt doorgeven.
- ⇨ Kan ik nu aan iemand vertellen hoe we zoal praten?

Bij wijze van besluit geeft de Leraar aan dat leerlingen in groepen moeten leren werken omdat ze, precies door met elkaar te praten, ideeën aan elkaar kunnen doorgeven en van elkaar kunnen leren.

Verdiepingsactiviteit

Materiaal

- ✓ Luisteroefening: fragment uit : 'Appartemensen' van Dimitri Leue ('De boer').

Opdracht 1 (in groepjes van drie)

De leerlingen luisteren eerst naar het verhaal (tot 2:56) . Daarna voorspellen ze het verdere verloop van het verhaal. Wie is Rosalie en waarom is ze weg? Vervolgens beluisteren ze het vervolg van het verhaal.

Korte klassikale nabespreking : hoe hebben de leerlingen het verhaal ervaren: prettig, verwarrend, grappig, vervelend… ?

Opdracht 2 (idem)

De leerlingen krijgen één grote tekening die bij het verhaal hoort en een aantal kleinere tekeningen. Ze overlopen hoe het verhaal liep en passen de kleine tekeningen in. Daarna bespreken ze de volgende vragen:

- ⇨ Op welke momenten verstaan Bonnie en de boer elkaar niet?
- ⇨ Hoe komt het dat de personages elkaar verkeerd verstaan?
- ⇨ Welke andere geluiden hoorden jullie? Werken die storend of helpen die juist om beter te begrijpen wat de boodschap is?

Opdracht 3 (idem)

In tekststroken schrijven de leerlingen wat Bonnie en de boer eigenlijk <u>echt</u> willen vertellen aan elkaar.

- ⇨ De leerlingen bespreken samen waarom het belangrijk is om naar elkaar te luisteren.
- ⇨ Welk verschil maakt het als je goed naar iemand luistert?
- ⇨ Kun je luisteren en tegelijk denken aan wat mensen aan het zeggen zijn?

 Bonnie
 Polluc

Materiaal

- ✓ APPARTEMENSEN, een verhaal met voetnoten in vier verdiepingen (recensie Pluizer)

auteur: Dimitri Leue - *illustrator:* Tom Schoonooghe

Dit kleine boekje met bijhorende luister-cd vertelt het verhaal van mensen die in hetzelfde appartementsgebouw 'Dahlia' wonen in de bloemenwijk. Eén van de bewoners is Bonnie. Zij ziet het als haar persoonlijke missie om iedereen in het appartementsgebouw te leren kennen en hen met elkaar in contact te brengen. Ze trekt haar stoute schoenen aan en gaat bij alle bewoners op bezoek. Zoals de titel van het boek reeds doet vermoeden, wonen er in het gebouw heel wat aparte mensen; hun diversiteit zorgt voor een vrolijke mix van figuren. Dimtitri Leue maakt in dit verhaal graag gebruik van grappige woordrijmpjes en geeft de meer geoefende lezer een duidelijke boodschap mee van respect en verdraagzaamheid. De teksten zijn eerder geschikt voor oudere lezers. De illustraties zijn van de hand van Tom Schoonooghe met tekeningen die een kinderlijke indruk geven. Naast het verhaal en de illustraties vind je op elke pagina eveneens gekleurde kaders terug waarin begrippen zoals het woord 'moslim' verder uitgelegd worden. Hierdoor oogt het geheel nogal druk en schept dit verwarring tijdens het lezen. Het is dan ook een mogelijkheid om het verhaal te lezen zonder de kaders te bekijken om zo de rode draad niet te verliezen. De luister-cd werd ingesproken door Britt van Marsenille, gekend als wrapper bij Ketnet, en is een absolute aanrader! Tevens hoor je onder andere de stemmen van Dimitri Leue en Tom Schoonooghe als bewoners van het gebouw. Op een zeer humoristische manier wordt het verhaal voor de luisteraar gebracht, waardoor deze meteen geboeid wordt en niet kan afgeleid worden door kadertjes met extra uitleg en tekeningen.

Basisles 2

Doelen De leerlingen worden zich bewust hoe belangrijk het is naar elkaar te luisteren. Succescriteria Ik kan luisteren naar wat andere kinderen zeggen. Daarna kan ik vertellen waarover ze het hadden.	Verdieping Focus ✓ Vertellen ✓ Aan de beurt zijn ✓ Onderbreken ✓ Verduidelijken ✓ leerlingen zijn aan het leren om goed naar iemand te luisteren

Materialen

- ✓ Babbelkaartjes (‘Sprekend leren’, p. 61)
- ✓ Woordkaartjes (‘Sprekend leren’, p. 74)
- ✓ Keukenwekker
- ✓ Track 6 uit ‘Appartemensen’
- ✓ Afbeelding voor fluisteropdracht (zie hieronder)

Klassikaal werk

Kringgesprek. De leerlingen halen woordkaartjes uit de doos. Ze leggen het woord uit, gebruiken het in een zin. De leerkracht legt uit dat al deze woorden te maken hebben met spraak en dat spreken en luisteren allebei belangrijke dingen zijn om te leren, zeker als je het goed wil doen op school.

Groepswerk

Materialen

- ✓ Babbelkaartjes PRAAT en LUISTER (voor elke deelnemer)
- ✓ Keukenwekker (‘Sprekend leren’, p. 69)

Klassikaal moment 2

De leerkracht verzamelt samen met de leerlingen de kenmerken van een goede luisteraar en een goede prater ('Sprekend leren', p. 70). Eventueel noteert hij deze kenmerken in een schema om ze in de klas op te hangen, zodat de leerlingen ze blijvend kunnen zien. Bovendien kan hij de antwoorden illustreren op basis van de input die leerlingen eerder over elkaar gaven, met name in het rode luik van het schema uit basisles 1.

Iemand die goed praat :	Iemand die goed luistert
✓ Iemand die iets uitlegt ✓ Iemand die goed op vragen antwoordt ✓ Iemand die naar je kijkt als hij iets vertelt ✓ …	✓ Iemand die je aankijkt ✓ Iemand die stil zit ✓ Iemand die laat zien dat hij geïnteresseerd is in wat je zegt ✓ Iemand die een vraag stelt en naar het antwoord luistert

Verdiepingsactiviteit

Opdracht 1

De leerlingen vormen een rij van 10. Ze krijgen de tekening nog niet te zien. Er wordt een zin doorgefluisterd. De laatste leerling tekent wat hij gehoord heeft. Is het juist? Wat liep er mis?

Herhaal deze activiteit, maar ditmaal kunnen de leerlingen de tekening wel zien (niet opgenomen in dit proefschrift wegens copyright).

- ✓ Waarom is het moeilijk?
- ✓ Was de eerste keer moeilijker of niet? Hoe komt dat denk je?
- ✓ Wanneer hebben de leerlingen dit al eens in het echte leven meegemaakt?

1. Op de witte uientoren staat een groene ster met acht punten en een vogel vliegt door de zon.

2. Onderaan de boom zit een droevige kat met een haar staart om de stam gewikkeld naar een appel op de grond te staren.

Opdracht 2

Materiaal

- ✓ Luisterfragment (track 6) uit Dimitri Leue (2009). *Appartemensen*. Tielt, Lannoo: 'Appartement 2A, Al Hawasi'

Groepswerk

De leerlingen luisteren naar het gesprek tussen Bonnie van 4A en Fatima van 2A. Vanwege het accent van Fatima en het gebruik van Marokkaanse woorden en namen, en een Nederlandse uitspraak met een accent moeten de leerlingen goed luisteren naar de inhoud. Bovendien kunnen sommige gebruiken voor deze klas vreemd of onbekend zijn.

In multiculturele klassen biedt deze info een extra element om met elkaar te praten en in discussie te gaan en de babbelregels te leren gebruiken.

De leerkracht laat de leerlingen na het beluisteren van het fragment de namen van de familieleden opnoemen. Zijn ze juist? Spreken ze deze juist uit met het accent op de correcte plaats? (Fatimà; Abdelassír; Mohamméd; Ràchida; Alì)

Voor de verwerking kan de leerkracht de leerlingen een aantal staakwoorden aanbieden. Eén leerling in de groep begint te vertellen waar het woord in het verhaal voorkomt (reconstructive van het verhaal) waarna hij zijn interpretatie ervan geeft of vertelt wat hij erover weet. Een volgende leerling mag aanvullen.

De leerlingen mogen aan elkaar om hulp vragen, vragen stellen, nagaan of iedereen akkoord gaat met het verloop van het verhaal of de uitleg die gegeven wordt.

Mogelijke thema's:

- schoenen uitdoen
- spelletje van de namen
- zwarte kubus
- begroeting en afsluiting : wat zegt Fatima?
- boerka en verstoppen
- Irak

Klassikale nabespreking op basis van de volgende vragen:

- ⇨ Werd de verteller soms onderbroken?
- ⇨ Hoe voelde hij zich daarbij?
- ⇨ Kan het nuttig zijn iemand te onderbreken?
- ⇨ Hoe kan je aan de andere laten weten dat je het goed vindt dat ze suggesties doen terwijl je aan het praten bent?
- ⇨ (Hoe kun je de vertelbeurten goed organiseren?)

Basisles 3

Doelen	Verdieping activiteit
• In groep beslissingen leren nemen. Succescriterium • De kinderen van onze groep kunnen luisteren • De kinderen van onze groep zeggen waarom ze iets denken • De kinderen van onze groep proberen samen een beslissing te nemen.	Focus op • Individuele beslissingen versus groepsbeslissingen

→ Succescriteria:

- ⇨ De kinderen van onze groep kunnen luisteren
- ⇨ De kinderen van onze groep zeggen waarom ze iets denken
- ⇨ De kinderen van onze groep proberen samen een beslissing te nemen.

Materialen

- ✓ Babbelkaartjes (‘Sprekend leren’ p. 61 of website) per kind, geknipt en gelamineerd: ‘ Wat denk jij?’ en ‘Waarom denk jij dat?’
- ✓ Woordkaartjes luisteren en spreken (‘Sprekend leren’ p. 74 of website) geknipt en gelamineerd.
- ✓ Presentatie workshops van http://www.graffitivzw.be/
- ✓ Strategiekaart voor groepsopdracht
- ✓ Foto’s
- ✓ 4 strookjes papier per groepslid + balpen
- ✓ 1 A4-blad per groep
- ✓ Kopieën van de gemaakte teksten per aantal deelnemers in de groep + enkele extra’s
- ✓ Tekengerei

Klassikaal werk 1

Voor deze les is een uitdagende context nodig : de leerlingen leren wat het betekent om om de beurt te spreken, naar elkaar te luisteren, elkaars mening te vragen en kritisch door te vragen.

We kiezen voor het samen uitzoeken van een workshop. Het materiaal is ontleend aan http://www.graffitivzw.be/

- De leerkracht vertelt de leerlingen dat hij wil weten welke workshop ze graag in hun klas zouden volgen. Er is echter een probleem: er is een aanbod van 5 workshops maar er kan maar één georganiseerd worden.

Toon (eventueel projectie op digibord) de verschillende mogelijkheden (cf. infra).

- De leerkracht vraagt wat de leerlingen van de voorstellen vinden. Ze bespreken wat 'fair' is en waarom.
- De leerkracht laat de leerlingen bij handopsteking kiezen voor een workshop. Hij legt uit uit dat stemmen tot verschillende resultaten kan leiden en turft de resultaten op het bord.
- De leerkracht bespreekt met de leerlingen wat een faire manier is om te beslissen welke workshop het wordt. Eventueel doet hij enkele suggesties (Voorbeelden, zie 'Sprekend leren', p. 76).

GRAFFITI Laat je verf en penselen maar thuis, want vandaag schilder je met een spuitbus. Durf jij het aan om met een echte graffiteur samen te werken? Eerst leert hij je de verschillende spuittechnieken, daarna ontwerp en spuit je je eigen muurschilderij! Oefen alvast je vingerspieren!
Komt aan bod :graffititechnieken met spuitbus, eigen graffiti-ontwerp uitwerken en spuiten paneel of muur met graffitiwerk.

DINER POEPSJIEK Vandaag ben je uitgenodigd op Diner poepsjiek! Je bereidt samen met je vrienden een lekker driegangenmenu. Tussen het kokerellen door train je de tafel te dekken als een opperbeste ober, onder het alziend oog van de strenge etiquettespecialist!
Komt aan bod : introductie in de tafeletiquette, snijtechnieken, goochelen met potten en pannen ,eenvoudig driegangenmenu om van te smullen.

KAST VOL POPPEN Word poppenmeester van het kleinste theater ter wereld: de poppenkast! Eerst maak je een gekke, grappige of griezelige pop. Daarna ga je met je pop op avontuur. Met een vleugje fantasie hier en een paar gekke geluiden daar breng je je pop tot leven.
Komt aan bod : poppentheatertechnieken: hand- en stemtraining, eenvoudige pop knutselen, verhaal uitwerken, pop-performance in groep

FOTOPOP Saaie kiekjes? Awoe! Daar zijn wij tegen! We leren je eenvoudige trucjes om te gekke foto's te nemen. Doen alsof je twee vrienden op je handen draagt, in je eigen schoen kruipt of een superheld bent: het is allemaal mogelijk.
Komt aan bod : basistechnieken fotografie: creëren van optische illusies, visuele mopjes uitwerken, knutselen, foto's nemen, foto's met optische illusies zonder digitale bewerking.

De leerkracht legt uit dat de klas in groepen verdeeld wordt en dat ze in die groepen nu een miniworkshop krijgen. De leerlingen mogen elkaars mening niet zomaar mogen goedkeuren of afwijzen. Ze moeten waarom-vragen stellen, redenen aangeven waarom ze een idee bruikbaar of niet vinden. De praat-en luisterwoorden die in basisles 2 werden geïntroduceerd, worden herhaald. De woordkaartjes worden getoond, ter herinnering.

Groepswerk (groepjes van 3)

De leerlingen duiden een verslaggever aan (voor zijn taakomschrijving zie 'Sprekend leren', p. 78)

Materialen

- ✓ Babbelkaartjes en Strategiekaart (cf. supra).

Opdracht

- Laat de groepjes 1 foto kiezen uit een aantal foto's. Dit kunnen foto's uit tijdschriften zijn maar ook foto's van activiteiten op school.

- Vraag hen in stilte te kijken naar de foto's en op de voor- en achterzijde een woord of korte zin te schrijven waaraan ze denken bij de foto. Nodig hen uit om zowel werkwoorden, bijvoeglijke naamwoorden als zelfstandige naamwoorden te gebruiken (10').
- Eén leerling legt een eerste woord op de tafel. De andere leerlingen leggen er een woord bij dat aansluit bij het eerste woord. Zo bouwen ze een 'tekst ' op door 12 woorden en/of zinnen onder elkaar te leggen. Ze bespreken daarbij of er woorden van plaats moeten veranderen. Ze mogen woorden ook aanpassen indien nodig (bv. vervoeging) of een woordje toevoegen of weglaten.
- (De leerlingen bespreken met elkaar de volgorde, Wijzigingen … en leggen aan elkaar uit waarom ze dat woord schreven, waarom ze een wijziging willen doorvoeren.)
- (Als de groep tot een consensus gekomen is, schrijven ze hun tekst op het A4-blad over.) → dit kan steekproefgewijs ook klassikaal besproken worden.
- Vervolgens draaien ze de woordstrookjes om: ze krijgen een (heel) andere tekst. Is deze beter/minder goed? Wat moet er gebeuren? → eerst kijkt elk groepslid naar de nieuwe woorden en zoekt zelf verbanden; vervolgens zegt elke leerling wat hij zou wijzigen en waarom; dan een nieuwe volgorde maken.

<u>Voorbeelden</u>

groot	man met rode short	blote voeten	fruit snoepen	gevaarlijk	opletten

gevaarlijk	dikke tenen	brommers	kiezen met slurf	druk	dikke olifant

druk brommers gevaarlijk dikke olifant groot honger kiezen met slurf fruit snoepen dikke tenen blote voeten man met rode short opletten	Druk Veel brommers Gevaarlijk Dikke olifant Groot Grote honger Kiest met lange slurf Snoept fruit Dikke olifantentenen Blote voeten Man met rode short : pas op!

Zesde leerjaar De Zevensprong Herfstwandeling oktober 2014
http://zesde.weebly.com/6b.html verkregen op 17/02/2015

konijn	wit	laarzen	zon	opgepeuzeld	grote ogen
dode vogel	skelet	bladeren	Hasnae	doodshoofd	ruggengraat

Bladeren	Bruine bladeren
Laarzen	Zwarte laarzen, bloemenlaarzen
Hasnae	Hasnae
Skelet	Vindt een skelet
Dode vogel	Een dode vogel?
Konijn	Een konijn?
Opgepeuzeld	Opgepeuzeld
Ruggengraat	Witte ruggengraat
Grote ogen	Grote oogkassen
zon	Ze zien de zon niet meer

De leerkracht kan de leerlingen ondersteunen door hen tips te geven of een extra opdracht. In bovenstaande opdracht ligt de nadruk om het toevoegen van adjectieven en preciezer te omschrijven.

Klassikaal moment 2

De resultaten van het Groepswerk worden klassikaal overlopen.

- De foto met de tekst wordt voorgelezen . De foto's zijn duidelijk zichtbaar voor alle leerlingen.
- De verslaggever licht de werkwijze toe en kan ook uitleggen waarom bepaalde keuzes gemaakt werden, veranderingen werden doorgevoerd.
- Vraag naar mogelijke problemen die de leerlingen tegenkwamen en hoe ze deze oplosten. = terugblik op het gesprek zelf.

Verdiepingsactiviteit

Focus : individuele beslissingen vs. groepsbeslissingen

- Voorlezen van de gemaakte tekst in de eigen groep
- Elke leerling van de groep illustreert de groepstekst.
- Daarna vergelijken de leerlingen hun illustraties met elkaar en bespreken ze welke assemblage van tekeningen het best bij hun tekst past. Deze komt in de klas of op gang te hangen voor het 'publiek'.

⇨ De leerlingen worden zich bewust dat hoe groepsbeslissingen worden gemaakt.

⇨ Het accent liggen op de vraag : ' Hoe kom je tot een gemeenschappelijke beslissing'?

	STAPPENPLAN
1.	Kies met je groepje een foto uit. Overleg!
2.	Elk groepslid neemt 8 strookjes papier.
3.	Elk groepslid schrijft op de voor- en achterzijde van zijn strookjes , één woord, enkele woorden of een korte zin ➔ waaraan denk je als je deze foto ziet? Wat zie je? Wat zou je kunnen ruiken, horen, voelen? Je werkt alleen en in stilte.
4.	Iemand van de groep legt een eerste strookje op het grote blad.
5.	Heb jij een strookje dat hierbij aansluit : leg het erbij. Zo ga je door tot alle strookjes op zijn. Bespreek samen hoe je de strookjes ordent tot een tekst. Leg uit aan elkaar waarom je voor deze volgorde kiest, waarom je vindt dat jouw strookje het best aansluit.
6.	Lees je tekst
7. 8.	Vul aan, schrap woorden, maak zinnen, vervoeg de werkwoorden, wissel woorden van plaats. Overleg steeds en leg aan elkaar uit waarom je iets wil veranderen . Je mag de wijzigingen op het blad bijschrijven.
9.	Lees je tekst nog eens. Iedereen tevreden? Schrijf je tekst over op het blad.
10.	Draai nu de kaartjes om: je krijgt en heel andere tekst. Wat vinden jullie ervan?

Basisles 4

Doelen	Verdieping activiteit
• We leren samen praten om dingen te rangschikken en ideeën aan elkaar door te geven Succescriterium • Ik kan zeggen welke dingen leerlingen leuk vinden die ikzelf ook leuk vind	Focus op • Individuele beslissingen versus groepsbeslissingen

Materialen

- ✓ Babbelkaartjes (boek p. 61 of website) per kind, geknipt en gelamineerd: ‘praat’ en ‘luister?’
- ✓ 3 grote hoepels
- ✓ 3 groepjes kaartjes
- ✓ 3 soorten voeding: snoep, groenten, frietjes
- ✓ 3 schoolvakken: wiskunde, taal, turnen.
- ✓ 3 sporten: dansen, voetballen, basketballen.
- ✓ Beamer , geluidsinstallatie, filmpje ‘ Partly Cloudy’
- ✓ Afbeelding wolken voor verdiepingsactiviteit 1 (cf. infra)
- ✓ Tabel voor verdiepingsactiviteit 2 (cf. infra)
- ✓ 1 A4-blad met venndiagrammen (cf. infra)

Klassikaal werk 1

Zie ‘Sprekend leren’, p. 80-81

Op het bord worden 3 cirkels getekend die elkaar gedeeltelijk overlappen. Elke cirkel wordt gemarkeerd met 3 soorten voeding: snoep, groenten, frietjes. Er wordt aan de leerlingen gevraagd wat ze het liefst eten. Misschien eten ze de 3 dingen wel graag. Enkele kinderen hangen hun naamkaartje in de cirkel van hun voorkeur.

De labels worden veranderd naar 3 schoolvakken: wiskunde, taal, turnen. Vervolgens wordt de activiteit herhaald.

De labels worden voor een derde keer veranderd naar 3 sporten: dansen, voetballen, basketballen. Idem.

De volgende vragen komen, verspreid over de drie keuzemomenten, aan bod (als de leerlingen hun naamkaartje ergens hebben gehangen, worden er vragen gesteld):

- Wat zou je doen als je twee dingen even graag eet/doet?

- Waar hang je je kaartje als je geen enkel van de 3 graag doet/eet?
- Heb je dezelfde voorkeur als iemand anders in de klas?
- Zijn de redenen dan hetzelfde?
- ⇨ Doel van de les duidelijk maken: We leren praten om dingen te rangschikken en ideetjes aan elkaar door te geven.

Groepswerk (groepjes van drie, met verslaggever)

Materiaal

- ✓ Babbelkaartjes

Opdrachten

- De leerlingen bekijken het filmpje Partly Cloudy https://www.youtube.com/watch?v=-a6Pe1ovKHg&list=PLfs4d4BcyOy6P0jaICLxHZGE3XOEloINQ
- Klassikale nabespreking: wat hebben de leerlingen het best onthouden? Wat was grappig zielig, stom …? Wat is de boodschap van het verhaal?
- De leerlingen praten nu in groepjes over hun beste vriend(in) en vullen de volgende zinnen aan:

  - Een vriend is iemand die …
  - Een vriend zal nooit …
  - Een vriend doet …

- De leerkracht vraag de leerlingen om al pratend minstens twee dingen te ontdekken die ze gemeen hebben in hun opvatting over wat een echte vriend is. Ze vragen elkaar ook argumenten, bv. ‘ Waarom is samen lachen met een vriend anders dan met iemand anders?’
- De verkregen informatie wordt gebruikt om in de ‘wij houden van …’-kaders te plaatsen.
- De verslaggever werkt het groepsblad verder af, in overleg met de rest van de groep.
- Bespreek met de leerlingen wat ze hebben ervaren tijdens dit gesprek
  - ⇨ Ik kan uitdrukken welke dingen (in en vriendschap) anderen leuk of belangrijk vinden die ikzelf ook leuk of belangrijk vind.

Klassikaal moment 2

De resultaten van het groepswerk worden klassikaal besproken.

Zie ‘Sprekend leren’, p. 81

Verdiepingsactiviteit 1

Focus : vragen stellen, ideeën en argumenten uitwisselen, de mening van een ander respecteren.

De leerlingen krijgen een 'wolk'. Ze schrijven er hun naam in. De leerkracht nodigt hen uit om in de wolk van iemand anders iets te tekenen wat bij die persoon hoort (eigenschap, iets dat hij/zij goed kan, kent. Iets moois, iets bijzonders, een typisch aspect van die persoon.) Leg de nadruk op dat het gaat om <u>positieve</u> elementen.

De eigenaar van de wolk moet de tekening goedkeuren, anders moet de ander er nog even aan doorwerken. De leerlingen kunnen ook (mondeling of schriftelijk) een verzoek indienen voor een bepaalde tekening. Wanneer er voldoende tekeningen op de wolk staan, keren de leerlingen terug naar hun groepjes.

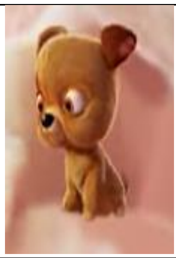
De leerlingen geven elkaar feedback : wat vinden ze mooi, ontroerend, onjuist , verrassend. Nodig hen ook uit om argumenten te geven bij hun uitspreken.

De leerkracht legt de nadruk op het respecteren van de gemaakte tekeningen.

- ⇨ De leerlingen geven aan wat ze mooi vinden of niet en kunnen hiervoor argumenten geven.
- ⇨ Het accent ligt op het respecteren van de visie van iemand anders én van de andere persoonlijkheid.

Verdiepingsactiviteit 2

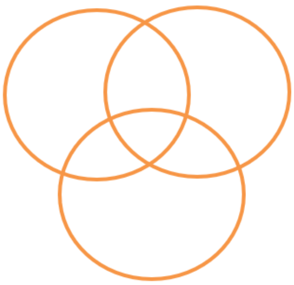
Focus : argumenten noemen om dingen te rangschikken. Hoe gebruik ik taal om logische verbanden aan te brengen en/of om te structureren?

Materiaal

- ✓ Babbelkaartjes ‘ik vind het goed, want’ en ‘ik vind het niet goed, want’.

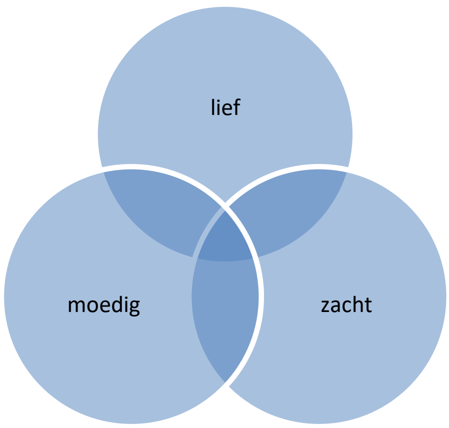
In hun groepjes van drie schrijven de leerlingen in drie kolommen individueel de eigenschappen op van de belangrijkste groepen uit de cartoon: de ooievaars, de baby’s en de wolken. Ze vergelijken en kiezen samen drie eigenschappen (fysiek of karakter) en verzamelen ze in een venndiagram.



De leerlingen plaatsen de verschillende personages in de venndiagrammen en bespreken waarom ze een personage al of niet in een overlapping zetten. Vervolgens vergelijken ze hun werk met dat van andere groepjes en bespreken onderling vragen als: Is dit een goede/ minder geslaagde keuze? Zou je ‘deze’ eigenschap ook aan je eigen venndiagram toevoegen? Waarom?

⇨ De leerlingen kunnen verwoorden waarom ze een personage bij een bepaald label ordenen.

Voorbeelden:



- Een kleine steenbok is wel lief maar niet zacht. Alle wolken zijn zacht en lief maar alleen de klunswolk is ook moedig omdat hij niet opgeeft.
- We hebben geen goede eigenschappen gezocht want we kunnen alleen de pechooievaar bij lief en moedig zetten.

De leerlingen kunnen ook bespreken dat bv. ‘ lief’ verschillende ladingen dekt.

Basisles 5

Doelen We proberen regels op te stellen voor onze groepsgesprekken Succescriterium • We hebben een reeks babbeldoosregels. We willen die allemaal gebruiken	Verdieping activiteit Focus op • de babbeldoosregels oefenen

Materialen

- ✓ tekeningen: speelgoedauto, voetbal, spel, vliegtuig, keuken, zwembad
- ✓ alle babbelkaartjes
- ✓ kopie uit Slegers, M. (2014)' Gezocht : normale ouders'. Alkmaar, Kluitman p. 7-13
- ✓ kopie achterflap van Slegers, M. (2014)' Gezocht : normale ouders'. Alkmaar, Kluitman
  - 3 deeltjes :
  - Kawipiwipi
  - Het hele kleine grootse avontuur
  - Een tijdloos avontuur

Dit is de sleutelles! De kinderen stellen zelf een aantal basisregels op voor groepsgesprekken. Ze baseren zich daarbij op hun ervaringen uit de vorige lessen. Ze spreken af om die tijdens het groepswerk te gebruiken. Dit zijn de **Babbeldoosregels**. Verwijs nog eens naar de inleiding, naar de Babbeldooslessen. Lessen 1 en 4 hebben de kinderen bewust gemaakt van spreken en luisteren als middelen om samen te denken. Ze hebben hen geholpen om argumenten te vinden, te formuleren, te aanhoren. Ze hebben hen geholpen om samen tot beslissingen te komen. De volgende regels moedigen leerlingen aan om gesprekken te voeren die op onderwijskundig vlak effectief zijn. We noemen het exploratieve gesprekken of sprekend leren. De voorgaande lessen en deze les stimuleren de leerlingen om deze regels zelf te ontdekken.

<u>Basisregels voor sprekend leren</u>

In deze vijfde les moeten de leerlingen (in hun eigen woorden) hun eigen regels opstellen, meer bepaald de basisregels voor exploratieve gesprekken. De leerkracht heeft hier een belangrijke sturende rol. Het is bijvoorbeeld erg belangrijk kinderen erop te drukken dat in groepsgesprekken alle meningen gewaardeerd worden en dat elke bewering moet gestaafd worden met argumenten. Wijs er ook op dat constructieve uitdaging en argumentatie deel uitmaken van exploratieve gesprekken.

Kinderen stellen vaak gespreksregels voor die in feite slaan op correct gedrag, bv.:

'Stil zijn in de gang en in de bibliotheek.'

‘Niet roepen in de klas.’

Dergelijke regels hebben hun nut, maar ze zorgen er weinig of niet voor dat er een exploratief gesprek op gang komt. Maak dit ook duidelijk aan de leerlingen. Vraag hen om over de regels te praten. Vraag hen ook om tijdens het gesprek hun eigen regels te volgen (die eventueel worden geprojecteerd of opgehangen in de klas). Wil je er zeker van zijn dat de leerlingen bepaalde regels zeker formuleren, leg ze dan situaties voor waarin die regels nodig zijn. Voorbeelden van gesprekken die in de klas hebben plaats gehad, werken het best: de leerlingen leggen dan onmiddellijk de link met hun eigen concrete ervaringen.

Hieronder vind je een voorbeeld van de Babbeldoosregels die een klas heeft opgesteld.

Onze Babbeldoosregels:

1. Werk samen – probeer overeen te komen

2. Praat en luister om beurten

3. Zeg wat je denkt

4. Vraag waarom

5. Denk samen over de ideeën van iedereen

6. Probeer het eens te worden met elkaar

Klassikaal werk 1

De leerlingen nemen tekeningen uit de doos. Er wordt telkens gevraagd of er regels zijn die daarbij horen.

De basisregels worden geïntroduceerd. Dit zijn regels die iedereen kent en iedereen moet volgen. Ze hebben te maken met veiligheid, eerlijkheid, hygiëne… Wat zou er gebeuren als mensen die regels niet zouden volgen? (eigen spelregels maken, handen niet wassen, 4 koffers meenemen op het vliegtuig…) Regels werken alleen als iedereen ze volgt.

Hoe leer je die? Sommige regels ken je gewoon, bv. niet eten in het zwembad, niet rechtop staan in de auto… Bij wie zijn er thuis ook regels of afspraken? Heb je ook regels die iets te maken hebben met praten?

- elkaar niet onderbreken
- niet praten met de mond vol
- niet roepen naar elkaar
- …

De leerkracht geeft een nieuwe context aan: we winnen allemaal wel eens graag een spel en doen dan vaak alles om te kunnen winnen, we zijn competitief. Maar dit werkt niet in een gesprekje, daar luistert en praat iedereen om beurten.

- Leg uit dat je samen gespreksregels voor je klas gaat opstellen. Licht toe dat het aspect competitie niet werkt bij groepsgesprekken. Het gaat om samenwerken, delen.
- Gebruik hiervoor de vragenlijst uit 'Sprekend leren', p. 87.
- Wat werkte goed in de groep terwijl jullie aan het praten waren?
- Wat liep niet zo goed?
- Wie kan goed spreken/luisteren?
- Hoe weet je dat?
- Welke basisregel zou je groep beter kunnen doen samenwerken?
- Je kunt ook gebruik maken van het uitgewerkte schema uit basislessen 1 en 2

Ik	Ik en jij
Waarom praten?	Wanneer is praten lastig?

Groepswerk (groepjes van 3)

Materiaal

- ✓ Zie 'Sprekend leren' p. 88, werkblad p. 91
- ✓ Elke groep krijgt het werkblad 5.
- ✓ Babbelkaartjes Ik praat, ik luister. Ik vind dit goed want… Ik vind dit niet goed, want…

Een leerling leest het kaderstukje. De leerlingen vullen het werkblad in na overleg. Daarna praat de hele groep erover, wisselt meningen en ideeën uit e.d. met als doel te beslissen of dit:

- een zinnige regel is voor groepsgesprekken (groen)
- geen zinnige regel is (rood)
- misschien zinnig is of niet, we twijfelen (oranje)

Indien nodig mogen de leerlingen hun babbelkaartjes gebruiken om de discussie meer structuur te geven. Geef de leerlingen voldoende ruimte om toe te lichten wat ze in hun groep geleerd hebben.

De ideeën van onze groep:

Namen:

Praten met KLEUREN:
GROEN - goede ideeën
ROOD - geen goede ideeën
ORANJE - niet zeker

Praten en BESLISSEN:
Kies vier groene ideeën
Wij kiezen de nummers

Verzin zelf twee goede ideeën.
Noteer ze in de cirkels.

1. Degene die alles opschrijft, beslist wat er gebeurt.

2. Iedereen moet doen wat de leider zegt.

3. Wie het eerst iets zegt, beslist wat we gaan doen.

4. We houden onze ideeën voor onszelf, zodat niemand ermee gaat lopen.

5. We proberen elk de beste te zijn in onze groep.

6. Wie niet goed kan meedoen, negeren we.

7. We praten en luisteren om de beurt.

8. Niemand mag van gedacht veranderen.

9. Iedereen praat zo luid als hij maar kan.

10. We proberen samen te werken en goed met elkaar om te kunnen.

11. We luisteren naar elkaar en denken na over elkaars ideeën.

12. We vragen waarom iemand denkt dat ...

13. We denken dat het een goed idee is om gedachten uit te wisselen.

14. We proberen het eens te worden met elkaar.

15. We nemen beslissingen waar de hele groep achter staat.

16. Praten is eigenlijk hetzelfde als luidop denken.

Klassikaal werk 2

De groepen vertellen welke regels zij belangrijk vinden. Die worden opgeschreven op het bord.

Ook de twee extra regels worden besproken. Worden ze toegevoegd op het bord? Hoe ben je tot datzelfde besluit gekomen? Heeft iedereen zijn zegje kunnen doen?

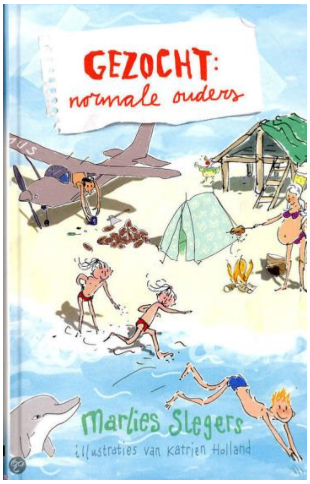
De leerkracht synthetiseert: “Dit noemen we vanaf nu de ‘babbelregels’. Je krijgt die op een kaartje en we gaan ze zoveel mogelijk gebruiken vanaf nu.”

⇨ Succescriteria gebruiken om te controleren of lesdoelen bereikt zijn.

Verdiepingsactiviteit 1

Focus : de babbeldoosregels oefenen.

De leerlingen zitten in hun groepjes.



- Vraag de leerlingen hoofdstuk O uit het boek ' Gezocht: normale ouders.' in stilte te lezen.
- Nodig hen uit om goed na te denken hoe ze de verschillende personages zouden tekenen.
- (De leerkracht kan er ook voor kiezen dit stukje voor te lezen. Bij deze aanpak vraagt hij de leerlingen om notities te nemen)
- Deel aan elk groepje een deel van de achterflap uit (het gaat om drie verhalen over deze knotsgekke familie :
  - Kawipiwipi
  - Het hele kleine grootse avontuur
  - Een tijdloos avontuur
- De groepen krijgen nu de opdracht om een voorflap te tekenen bij hun verhaal. Ze moeten hierbij de babbelregels gebruiken :
  - Ze selecteren herkenbare en tekenbare eigenschappen van de personages
  - Ze selecteren gegevens uit de korte inhoud van de verschillende verhalen (1 verhaal per groepje)
  - Ze beslissen samen hoe ze tot een uitnodigende kaft komen
- Maak de leerlingen attent op het gebruiken van de babbelregels.
- De leerlingen presenteren hun ontwerp en leggen uit hoe ze de babbelregels gebruikten om tot een beslissing te komen en om samen te werken. Je kan dit in de volledige groep doen of je kan groepjes met hetzelfde verhaal elkaars ontwerp laten vergelijken en bespreken.
  - ⇨ Welke regels vonden ze nuttig?
  - ⇨ Welke vonden ze moeilijk om te gebruiken?
  - ⇨ Welke vonden ze moeilijk om te respecteren?
  - ⇨ Denken ze dat ze beter werk afleverden als groep dan wanneer ze dit alleen hadden gedaan?

Verdiepingsactiviteit 2

Focus: de Babbeldoosregels oefenen

Geef elke groep een kopie van de Babbelkaartjes 'Ik ga akkoord want' en 'Ik ga niet akkoord want'.

Klassikaal werk

Activiteit A

Kies een onderwerp waarover gepraat wordt op school. Leg de kinderen een stelling voor of een vraag, bv. ‘Voetballen op de speelplaats moet verboden worden’, ‘Wie afval op de grond gooit, moet de speelplaats schoonvegen na schooltijd’ of ‘We willen cola in de drankautomaat’.

Activiteit B

Kies een van de volgende stellingen, vergroot ze, projecteer ze en leg uit wat ze wil zeggen:

- Kinderen zouden te voet naar school moeten komen.
- We zouden moeten leren typen, niet schrijven.
- We zouden geen wiskunde mogen leren, maar wel leren rekenen met een rekenmachine.
- Iedereen moet wekelijks evenveel zakgeld krijgen van de overheid.
- Kinderen mogen zelf beslissen hoe laat ze naar bed gaan.
- Kinderen mogen snoep of chocola krijgen als middageten.
- Kinderen moeten leren zwemmen.
- Kinderen moeten een schooluniform dragen.
- De schoolvakanties moeten elk een week langer zijn.

Groepswerk

Elke leerling neemt om de beurt een ‘Ik ga akkoord want’- kaartje, leest de stelling waarop het kaartje betrekking heeft en vult aan met een argument, bv. ‘Kinderen moeten leren zwemmen – Ik ga akkoord want anders zouden er aan zee veel kinderen verdrinken.’ De kinderen bespreken het argument: is het zinnig of niet? De kinderen gebruiken hun ‘Ik ga niet akkoord want’-kaartje om tegenargumenten te geven.

De groepen bespreken op die manier alle stellingen en komen telkens tot een groepskeuze: akkoord of niet akkoord. Als dat gebeurd is, brengt elke groep verslag uit: welke beslissing heb je genomen en waarom. Ga na in welke mate de basisregels bij dit proces hebben bijgedragen. Welke regels moeten er worden aangepast? Hoe doe je dat?

Geef de leerlingen voldoende ruimte om te expliciteren wat ze in hun groep geleerd hebben!

1.2 Poster ground rules (Dutch)

De poster met de babbelregels werd na de basislessen in het groot en op een duidelijk zichtbare plaats opgehangen in de klas, als referentiepunt voor Leerkracht en leerlingen.

Onze basisregels voor groepswerk

1. We moedigen elkaar aan om te spreken.
   ('Wat denk jij?')
2. We hebben respect voor elk idee en elke mening.
3. We kijken en luisteren naar elkaar.
4. We geven uitleg en vragen argumenten.
   ('Ik denk dat..., want...', 'Waarom denk je dat?')
5. We leggen uit waarom we het eens of niet eens zijn met iemand.
   ('Ik ga akkoord, want ...', 'Ik vind het niet goed, want ...')
6. We werken samen aan een besluit dat iedereen goed vindt.

1.3 Talking prompts (Dutch)

Ik vind het goed want …	Ik vind het goed want …
Ik vind het niet goed want …	Ik vind het niet goed want …
Waarom denk je dat?	Waarom denk je dat?
Wat denk je?	Wat denk je?

1.4 Assessment forms (Dutch)

Deze evaluatiefiches werden tijdens het quasi-experiment gebruikt door de leerlingen. Ze reflecteerden hiermee op individueel niveau en op groepsniveau op de groepsgesprekken zelf.

Groep: Datum:

	++	+	=	-	--
Alle leerlingen mochten hun mening zeggen of hun idee vertellen (‘Wat denk jij?’)					
Als iemand iets zei, werd er naar een argument gevraagd (‘Waarom zeg je dat?’)					
Alle leerlingen hebben actief geluisterd naar elkaar					
We hebben aanvaard dat niet iedereen het eens is met elkaar (‘Ik vind het goed want’, ‘ik vind het niet goed want’)					
We hebben samen gewerkt aan een besluit dat iedereen goed vindt					

De groepsleden

Naam:	++	+	=	-	--
Ik heb actief geluisterd					
Ik heb wat-vragen gesteld					
Ik heb waarom-vragen gesteld					
Wat ik zei sloot aan bij wat eerder gezegd was					
Ik heb anderen aangemoedigd om iets te zeggen					
Ik heb gerespecteerd wat iemand anders zei					
Ik heb meegewerkt aan een besluit dat iedereen goed vindt					

2. Teacher's log: standard form

WEEK + Thema	Les in de loop van de week. Noteer datum, lesonderwerp, gehanteerde basisregels en eventueel een opmerking met tip of een positieve ervaring	Les in de loop van de week. Noteer datum, lesonderwerp, gehanteerde basisregels en eventueel een opmerking met tip of een positieve ervaring	Les in de loop van de week. Noteer datum, lesonderwerp, gehanteerde basisregels en eventueel een opmerking met tip of een positieve ervaring
Week van ……………………… Inleidende les explorative gesprekken 1	Datum:	Datum :	Datum :
	Onderwerp:	Onderwerp:	Onderwerp:
	Besproken basisregel (omcirkel): 1 2 3 4 5 6	Besproken basisregel (omcirkel): 1 2 3 4 5 6	Besproken basisregel (omcirkel): 1 2 3 4 5 6
	Opmerking(en):	Opmerking(en):	Opmerking(en):
Week van …………………… Inleidende les exploratieve gesprekken 2-3	Datum:	Datum :	Datum :
	Onderwerp:	Onderwerp:	Onderwerp:
	Besproken basisregel (omcirkel): 1 2 3 4 5 6	Besproken basisregel (omcirkel): 1 2 3 4 5 6	Besproken basisregel (omcirkel): 1 2 3 4 5 6
	Opmerking(en):	Opmerking(en):	Opmerking(en):

3. Linguistic markers of exploratory talk

This list includes all unique key words in context and the category they belong in. Word groups or synonyms expressing the same meaning, as in the code book NVivo (Appendix 5), have not been included.

LM1 als
LM1 als dan
LM1 daarom
LM1 daarvoor
LM1 doordat
LM1 dan
LM1 en
LM1 en niet
LM1 omdat
LM1 ook
LM1 want
LM2 akkoord
LM2 eens
LM2 gelijk hebben
LM2 goed
LM2 goed idee
LM2 ja
LM2 oké
LM3 wat denk jij
LM3 wat is
LM3 wat vind
LM3 weet jij
LM4 waarom
LM4 waarom denk
LM5 besluit
LM5 dat klopt
LM5 dus
LM5 en dus
LM5 schrijven op
LM6 dat wil zeggen
LM6 denk ik
LM6 denken
LM6 maar
LM6 toch
LM6 vind ik
LM6 vind jij
LM6 volgens
LM6 wacht
LM6 zou

4. Checklist for exploratory lessons

This checklist was given to the teachers at the start of the experiment. They were asked to use it as guidelines for the construction of lessons or parts of a lesson which would focus on problem solving in groups.

	oké	aanpassen
Het groepswerk gaat over een concrete vraag, een concreet probleem dat de leerlingen enkel door samenwerking kunnen oplossen en waardoor ze samen kennis opbouwen.		
Het groepswerk heeft voor Leerkracht én leerlingen een duidelijk, concreet doel. *		
Het groepsgesprek is niet vrijblijvend maar is gericht op een (aanzet tot) oplossing. *		
De leerlingen worden uitgedaagd om een hypothese te vormen en daarvoor argumenten te formuleren.		
Het groepswerk daagt de leerlingen uit om via argumentatie tot een consensus te komen.		
De opdracht legt de focus op samen spreken en overleggen. Ze bevat een minimum aan schrijfwerk. (Gebruik eventueel sjablonen, kaartjes met begrippen…)*		
De leerlingen krijgen een duidelijke mondelinge instructie met expliciete verwijzingen naar de het gebruik van één of meerdere basisregels.		
De leerlingen beschikken ook over een schriftelijke instructie en een eventueel stappenplan.		
De materialen nodig voor het groepswerk zijn functioneel en leiden de leerlingen niet af van de taak.		
In deze les worden babbelkaartjes gebruikt en/of de basisregels zijn zichtbaar aanwezig in het lokaal.		
Er is tijd voorzien voor een vorm van nabespreking (individueel , in groep, klassikaal), met expliciete verwijzing naar de babbelregels		

5. Code book NVivo: queries key words in context

This codebook includes word groups expressing exactly the same meaning or synonyms of key words, e.g. 'eens zijn met' = 'akkoord'. For statistical purposes these were brought together and counted as one as (see appendix 3 and 6.2). Queries which were not found in any conversation were left out of the statistical dataset.

Queries

Name	Created On
LM1 als	16/01/2015 10:30
LM1 als dan	21/12/2015 17:36
LM1 daarom	16/01/2015 10:29
LM1 daarvoor	16/01/2015 10:29
LM1 dan	21/05/2015 13:08
LM1 doordat	16/01/2015 10:17
LM1 en	16/01/2015 10:28
LM1 en niet	16/01/2015 10:23
LM1 of	28/12/2015 17:02
LM1 omdat	16/01/2015 10:17
LM1 ook	9/04/2015 16:57
LM1 want	16/01/2015 10:16
LM2 akkoord	16/01/2015 19:52
LM2 eens zijn met	8/04/2016 13:15
LM2 gelijk hebben	27/04/2016 15:45
LM2 goed	16/01/2015 19:54
LM2 goed+idee	28/05/2015 13:47
LM2 ja	27/04/2016 11:19
LM2 OK	28/05/2015 13:48

Queries

Name	Created On
LM3 Wat denk	9/04/2015 17:01
LM3 Wat is	16/01/2015 10:48
LM3 Wat vind	23/12/2015 12:36
LM3 Wat vinden	27/04/2016 9:48
LM3 weet jij	21/12/2015 17:35
LM4 Waarom	16/01/2015 19:56
LM4 Waarom denk	9/04/2015 17:02
LM4 Wat denkt gij	3/05/2016 11:39
LM5 beslis	23/12/2015 12:00
LM5 besluit	28/05/2015 13:52
LM5 conclusie	28/05/2015 13:51
LM5 dat klopt	23/12/2015 12:34
LM5 dus	9/04/2015 17:25

Queries

Name	Created On
LM6 dat wil zeggen	16/01/2015 10:49
LM6 denk ik	14/05/2015 18:57
LM6 denken we	23/12/2015 12:35
LM6 geloof ik	21/12/2015 17:35
LM6 ik denk	14/05/2015 18:58
LM6 ik geloof	21/12/2015 17:33
LM6 ik vind	28/05/2015 13:48
LM6 ik weet	21/12/2015 17:34
LM6 maar	16/01/2015 10:18
LM6 maar dan	21/05/2015 13:08
LM6 toch	10/04/2015 10:04
LM6 vind ik	28/05/2015 13:50
LM6 vind+jij	28/05/2015 13:53
LM6 vindt ik	23/12/2015 12:51
LM6 volgens	23/12/2015 12:35
LM6 wacht	11/05/2016 15:14
LM6 zou	21/05/2015 13:11

6. Code book SPSS

6.1 Statistical analyses at group level

a. Main dataset

Nr leerling	1-198
school	1-5
klas	1-11
EG1CG0	Target groe*p* = 1; control groe*p* = 2
groep	1-65 groepen van drie leerlingen elk
gender	Binnen elk groepje: 1 = VVV; 2=VVM; 3=MMM; 4= MMV ; 5= VVMM
GOK	GOK indicator binnen elk groepje: 1 = ja ja ja of ja ja nee 2 = ja nee nee of nee nee nee
LVSspelling	Resultaten leerlingvolgsysteem spelling per groepje: 1 = homogeen hoog 2 = homogeen laag 3 = heterogeen
LVSwiskunde	Resultaten leerlingvolgsysteem spelling per groepje: 1 = homogeen hoog 2 = homogeen laag 3 = heterogeen
INSLIJPEN	Mate waarin de techniek van exploratieve gesprekken is ingeslepen tijdens het experiment: 1 = sterk (min. 2x per week) 0 = zwak (1xweek of minder)
gefilmd110	1 = groepje werd gefilmd 0 = groepje werd niet gefilmd
NULRPMgroep	Resultaten nulmeting Raven's Progressive Matrices per groepje
EFFRPMGROEP	Resultaten effectmeting Raven's Progressive Matrices per groepje
scoreverschilgroepen	Scoreverschil nul- en effectmeting Raven's Progressive Matrices per groepje
NULnc	Aantal linguïstische markers van exploratieve gesprekken gebruikt tijdens de nulmeting – discussie over een niet-leerstofgebonden onderwerp
NULnc-A	Idem, in procenten vergeleken met de woordlengte van het gesprek
TotaalArgNULnc2	Aantal argumenten gebruikt tijdens het gesprek
EFFnc	Aantal linguïstische markers van exploratieve gesprekken gebruikt tijdens de effectmeting – discussie over een niet-leerstofgebonden onderwerp
EFFnc-A	Idem, in procenten vergeleken met de woordlengte van het gesprek

ArgEFFnoc	Aantal argumenten gebruikt tijdens het gesprek
NULRaven	Aantal linguïstische markers van exploratieve gesprekken gebruikt tijdens de nulmeting – Raven's Progressive Matrices
NULRaven-A	Idem, in procenten vergeleken met de woordlengte van het gesprek
ArgEFFRaven	Aantal argumenten gebruikt tijdens het gesprek
EFFRaven	Aantal linguïstische markers van exploratieve gesprekken gebruikt tijdens de nulmeting – Raven's Progressive Matrices
EFFRaven-A	Idem, in procenten vergeleken met de woordlengte van het gesprek
ArgEFFRAven2	Aantal argumenten gebruikt tijdens het gesprek
verschil-LM-RPM	Verschil aantal linguïstische markers nul- vs. effectmeting - Raven's Progressive Matrices in absolute cijfers
verschil-LM-nc	Verschil aantal linguïstische markers nul- vs. effectmeting – groepsdiscussie - in absolute cijfers
verschil-ARG-nc	Verschil aantal argumenten nul- vs. effectmeting - groepsdiscussie
verschil-ARG-RPM	Verschil aantal argumenten nul- vs. effectmeting - Raven's Progressive Matrices
verschil-LM-RPM-procenten	Verschil aantal linguïstische markers nul- vs. effectmeting - Raven's Progressive Matrices – percentage van gesprek
verschil-LM-nc-procenten	Verschil aantal linguïstische markers nul- vs. effectmeting – groepsdiscussie – percentage van gesprek
Z-scoreverschilgroepen	Gestandaardiseerd verschil groepsscores Raven's Progressive Matrices
Z-verschilLM-RPM	Gestandaardiseerd verschil aantal linguïstische markers Raven's Progressive Matrices
Z-verschilLM-nc	Gestandaardiseerd verschil aantal linguïstische markers groepsdiscussie

b. Dataset long utterances

nummer	Nummer leerling
groep	Groepsnummer
NULL0 EFF1	Pre- vs. post-test
CG0 EG1	Controlegroep vs. experimenteergroep
NULncd tekens	Aantal karakters per beurt per leerling niet-curriculumgebonden discussie, pre-test
NULncd100	Aantal spreekbeurten met meer dan 100 karakters
NULncd100prop	Idem, proportioneel op basis van de volledige gesprekslengte
NULncd70	Aantal spreekbeurten met meer dan 70 karakters
NULncd70prop	Idem, proportioneel op basis van de volledige gesprekslengte
NULncd115	Aantal spreekbeurten met meer dan 115 karakters
NULncd115prop	Idem, proportioneel op basis van de volledige gesprekslengte
NULrpm tekens	Aantal karakters per beurt per leerling probleemoplossende discussie
NULrpm100	Aantal spreekbeurten met meer dan 100 karakters, pre-test
NULrpm100prop	Idem, proportioneel op basis van de volledige gesprekslengte
NULrpm70	Aantal spreekbeurten met meer dan 70 karakters
NULrpm70prop	Idem, proportioneel op basis van de volledige gesprekslengte
NULrpm115	Aantal spreekbeurten met meer dan 115 karakters
NULrpm115prop	Idem, proportioneel op basis van de volledige gesprekslengte
EFFncd tekens	Aantal karakters per beurt per leerling niet-curriculumgebonden discussie, post-test
EFFncd100	Aantal spreekbeurten met meer dan 100 karakters
EFFncd100prop	Idem, proportioneel op basis van de volledige gesprekslengte
EFFncd70	Aantal spreekbeurten met meer dan 70 karakters
EFFncd70prop	Idem, proportioneel op basis van de volledige gesprekslengte
EFFncd115	Aantal spreekbeurten met meer dan 115 karakters
EFFncd115prop	Idem, proportioneel op basis van de volledige gesprekslengte
EFFrpm tekens	Aantal karakters per beurt per leerling probleemoplossende discussie, post-test
EFFrpm100	Aantal spreekbeurten met meer dan 100 karakters, pre-test
EFFrpm100prop	Idem, proportioneel op basis van de volledige gesprekslengte
EFFrpm70	Aantal spreekbeurten met meer dan 70 karakters
EFFrpm70prop	Idem, proportioneel op basis van de volledige gesprekslengte
EFFrpm115	Aantal spreekbeurten met meer dan 115 karakters
EFFrpm115prop	Idem, proportioneel op basis van de volledige gesprekslengte
EFFrpmd115prop	Aantal karakters per beurt per leerling probleemoplossende discussie

c. Dataset turn-taking

Nr	leerlingnummer
groep	Groepsnummer
NULL0 EFF1	Pre- vs. post-test
CG0 EG1	Controlegroep vs. experimenteergroep
NULnc spreekbeurten	Totaal aantal spreekbeurten niet-leerstofgebonden discussie, pre-test
NULnc speling12	Symmetrisch (0) vs. assymmetrisch gesprek (1)
NULnc2045	Interactive dominantie/recessie nee (0) vs. ja (1)
NULrpm spreekbeurten	Totaal aantal spreekbeurten probleemoplossend gesprek, pre-test
NULrpm speling12	Symmetrisch (0) vs. assymmetrisch gesprek (1)
NULrpm2045	Interactive dominantie/recessie nee (0) vs. ja (1)
EFFnc spreekbeurten	Totaal aantal spreekbeurten niet-leerstofgebonden discussie, post-test
EFFnc speling12	Symmetrisch (0) vs. assymmetrisch gesprek (1)
EFFnc2045	Interactive dominantie/recessie nee (0) vs. ja (1)
EFFrpm spreekbeurten	Totaal aantal spreekbeurten probleemoplossend gesprek, post-test
EFFrpm speling12	Symmetrisch (0) vs. assymmetrisch gesprek (1)
EFFrpm2045	Interactive dominantie/recessie nee (0) vs. ja (1)

6.2 Statistical analyses at individual level

Meetmoment	Moment van meting pre-/post-test, niet-curriculaire discussie/probleemoplossende toets
NULL0EFF1	Pre- vs. post-test
CG0EG1	Controlegroep vs. experimenteergroep
Nr	Uniek nummer leerling
school	Schoolcodenummer
klas	Klascodenummer
groepnr	Nummer van het groepje waartoe de leerling behoort
leeftijd	Leeftijd leerling
GOK	Leerlingen heeft geen (0) / wel (1- GOK-indicatoren
M1V2	Meisje (2) of jongen (1)
AVI	AVI-niveau
LVS SP	Spellingniveau hoog (1) of laag (0)
LVS WI	Wiskundeniveau hoog (1) of laag (0)
NUL RPM groep	Score Raven's op groepsniveau, pre-test
NUL RPM ind	Score Raven's op individueelniveau, pre-test
EFF RPM GROEP	Score Raven's op groepsniveau, pre-test
EFF RPM IND	Score Raven's op individueelniveau, pre-test
scoreverschil groepen	Scoreverschil Raven's pre-/post-test groepen
scoreverschil indvid	Scoreverschil Raven's pre-/post-test individuele leerlingen
woordengesprek	Aantal woorden per conversatie
spreekbeurten	Aantal spreekbeurten per leerling
woordenzin	Aantal woorden per zin per conversatie
LM1 als	Talige markers categorie 1
LM1 als dan	
LM1 daarom	
LM1 daarvoor	
LM1 doordat	
LM1 dan	
LM1 en	
LM1 en niet	
LM1 omdat	
LM1 ook	
LM1 want	
SOM LM1	Totaal aantal talige markers categorie 1
LM2 akkoord	Talige markers categorie 2
LM2 eens	
LM2 gelijk hebben	
LM2 goed	
LM2 goed idee	
LM2 ja	
LM2 oké	
SOM LM2	Totaal aantal talige markers categorie
LM3 wat denk jij	Talige markers categorie 3

LM3 wat is	
LM3 wat vind	
LM3 weet jij	
SOM LM3	Totaal aantal talige markers categorie 3
LM4 waarom	Talige markers categorie 4
LM4 waarom denk	
SOM LM4	Totaal aantal talige markers categorie 4
LM5 besluit	Talige markers categorie 5
LM5 dat klopt	
LM5 dus	
LM5 en dus	
LM5 schrijven op	
SOM LM5	Totaal aantal talige markers categorie 5
LM6 dat wil zeggen	Talige markers categorie 6
LM6 denk ik	
LM6 denken	
LM6 maar	
LM6 toch	
LM6 vind ik	
LM6 vindt jij	
LM6 volgens	
LM6 wacht	
LM6 zou	
SOM LM	Totaal aantal talige markers categorie 6
TOTAAL LM	Totaal aantal talige markers alle categorieën
Proportioneelgebruikwoorden	Berekening talige markers, proportioneel
FXPRED 1	Linear Mixed Models - berekeningen
PRED 1	
FXPRED 2	
PRED 2	
RPMind	
FXPRED 3	
PRED 3	
FXPRED 4	
PRED 4	
FXPRED 5	
PRED 5	
FXPRED 6	
PRED 6	
FXPRED 7	
PRED 7	
FXPRED 8	
PRED 8	
FXPRED 9	
PRED 9	
FXPRED 10	

PRED 10	
FXPRED 11	
PRED 11	
FXPRED 12	
PRED 12	
FXPRED 13	
PRED 13	
FXPRED 14	
PRED 14	
FXPRED 15	
PRED 15	