Mathematical Thinking, Learning and Performance:

Reed, H.C.

2014

2

Preschoolers' causal reasoning during shared picture-book storytelling: A cross-case comparison descriptive study

Helen C. Reed, Petra P. M. Hurks, Paul A. Kirschner and Jelle Jolles

Submitted

ABSTRACT

This study investigates how shared picture-book storytelling within a peer-group setting could stimulate causal reasoning in children aged 4½ to 6 years. Twenty-eight children from preschool classes of three schools were allocated to one of six groups (4-5 children per group). Each group participated in six storytelling sessions over a period of two weeks. During these sessions, the children freely generated stories from pictorial stimuli in two picture-books. Storytelling discourse was analysed in the groups that showed the lowest and the highest pre-to-post-intervention improvement on a series of causal reasoning tasks. In the most-improving group, discourse was distinguished by detailed interpretations of perceptual features, causal explanations and explicit justifications of statements. The least-improving group was distinguished by 'superficial' talk (i.e., labelling perceptual features, simple inferences, uncritical acceptance of statements and disagreements). These types of discourse could be related to time spent on storytelling. The findings generate hypotheses for future research on stimulating causal reasoning in early childhood education.

Keywords: preschoolers; causal reasoning; narratives; peer interaction; picture-books

INTRODUCTION

Causal reasoning can be broadly defined as the ability to understand relations with a cause and an effect. Children's capacity for causal reasoning is fundamental to understanding the dynamic physical and social environments they confront every day (Hickling & Wellman, 2001; Schlottman, Allen, Linderoth, & Hesketh, 2002; Wellman & Liu, 2007). It allows them to understand past events, predict future outcomes and influence what happens in the world (Goswami, 2008). Later on, children need this capacity to be successful at school, where much of the formal thinking underlying school subjects - including mathematics, natural sciences and reading comprehension - is causal in nature (Devlin, 2000; Kuhn & Pearsall, 2000; Van den Broek et al., 2005).

This study investigates how causal reasoning could be stimulated in preschool children (aged 4½-6 years) through shared picture-book storytelling within a peer-group setting. We distinguish specific types of causal reasoning (Goswami, 2008). In *causal inference*, a conclusion is drawn about a connection between an antecedent and an outcome. *Causal explanations* describe mechanisms that act on antecedents to produce particular outcomes. *Causal predictions* are made about the outcome of an event, assuming a particular causal mechanism. *Counterfactual thinking* imagines how outcomes would have been different if causal antecedents were different.

Causal reasoning in young children

Young children already possess considerable causal reasoning skills. Between two and four years of age, children can understand that causal agents can bring about transformations in objects (e.g., Das Gupta & Bryant, 1989), seek and provide causal explanations for natural events, human activity and psychological states (e.g., Hickling & Wellman, 2001; Legare, Gelman, & Wellman, 2010), make causal predictions about future events (e.g., Bonawitz et al., 2010), start reasoning about counterfactuals (e.g., Beck, Robinson, Carroll, & Apperly, 2006), and infer causal relations from patterns in environmental variation (e.g., Gopnik et al., 2004). Although more complex (e.g., multivariable) causal reasoning abilities continue to develop up to 6 or 7 years or older (Kuhn & Pearsall, 2000; Lagattuta & Wellman, 2001), the ability to think about causal relations appears quite robust by the time children enter early childhood education.

There are, however, important individual differences in young children's causal reasoning skills, stemming from differences in knowledge and/or experience. Both are important preconditions for competent causal reasoning: young children use prior knowledge of causal structures and mechanisms to learn new causal relations and to make causal inferences and predictions (Gopnik et al., 2004; Schlottman, 2001). Considerable practice and experience are required before children learn to integrate various components of causal reasoning and the knowledge of causal structures that they acquire (Kuhn & Pease, 2008; Schlottman, 2001). A primary source of differences in knowledge and experience is variability in home context. Specifically, the frequency and type of talk engaged in between parents and children impacts children's understanding of the causes and consequences of everyday events and human behaviour (Wellman & Liu, 2007). Consequently, some children in preschool may lack much causal knowledge and/or experience in causal reasoning.

Causal reasoning in narrative construction

A fundamental way in which children apply causal reasoning is in constructing narratives - alone or in conversation with parents and peers - that help them derive meaning from their experiences (Paris & Paris, 2003; Schick & Melzi, 2010). Narratives are stories that construct relations between characters, actions, events, motivations and emotions, based on contextual clues and prior knowledge (Paris & Paris, 2003), and typically involve several types of causality. *Physical causality* refers to interactions between physical entities, as when an object transforms a second object from one state to another (e.g., a moving object bumping into a stationary one and setting it in motion), while *psychological causality* refers to how the behaviour of characters is related to mental states (e.g., running away because you are frightened) (Schlottman et al., 2002). In narratives, these types of causality often act in combination, as when behaviour and mental states are related to physical events (e.g., someone who is trapped becoming upset and crying). In this study, we refer to this as *situational causality*, reflecting the view that narratives involve mental representations of situations and that these representations include causal relationships between several story elements (Van den Broek et al., 2005; Zwaan & Radvansky, 1998).

Through constructing narratives, children learn to form coherent and meaningful mental representations embodying several knowledge structures, including concepts about causal relations between various story elements that underlie the order of events (Lever & Sénéchal, 2011; Paris & Paris, 2003). For example, when children construct narratives that make sense of pictorial stimuli, they make inferences, form predications and provide explanations that enable them to form coherent accounts of causally-sequenced plots (Paris & Paris, 2003). In a study by O'Neill, Pearce and Pick (2004), preschoolers' showed this kind of reasoning in narrating a picture-book story that depicted different mental states in the context of characters interacting (e.g., seeing a frog in her salad causes a diner to react with surprise and anger).

By around 5-years of age, children can produce well-structured, coherent and cohesive narratives (Schick & Melzi, 2010). However, there are considerable variations in narrative abilities that are associated with cultural and ethnic background, social or economic class and home context (Lever & Sénéchal, 2011; Peterson, Jesso, & McCabe, 1999; Schick & Melzi, 2010). These variations appear amenable to intervention. Peterson et al. (1999) helped mothers stimulate disadvantaged preschoolers' narrative skills by spending more time in narrative conversation, asking more open-ended and context-eliciting questions, and encouraging longer narratives. One year post-intervention, these children showed overall improvements in narrative skill, including more descriptions of when and where events took place. Hayward and Schneider (2000) taught narrative skills to preschoolers with language impairments through a number of activities, including sorting and sequencing elements, identifying missing elements, and reformulating scrambled stories. Post-intervention, children included more story information and produced more complex stories when generating narratives. In a study by Spencer and Slocum (2010), preschoolers with narrative language delays were taught narrative structure in a story-retell context using picture sequences and icons. Children made substantial gains in narrative retelling and personal story generation. Thus, interventions in which young children are encouraged to engage in narrative activities appear to stimulate their narrative abilities.

Reasoning through peer interaction

A second learning perspective relevant to the present study comes from sociocultural research: children's learning and development are significantly shaped by social and communicative interactions with peers (Howe & Mercer, 2010). Under appropriate conditions, peer interaction during learning tasks can stimulate children to think aloud, share knowledge, explore and express different views, engage critically and constructively with each other's ideas, and elaborate or defend their own points of view - a type of interaction called 'exploratory talk' (Howe & Mercer, 2010; Mercer & Howe, 2012). Importantly, such conditions include the absence of directive control by the teacher, whose role is to create conditions under which participants engage openly and equitably with each other's views (Mercer & Howe, 2012).

Although previous research has largely focused on older children (≥ 8 years) and adolescents, preschoolers also justify opinions, suggest alternatives, and reach compromises during free play with siblings or peers (Ehrlich, 2011; Howe & McWilliam, 2001). When preschoolers share stories with peers, they take on different roles than when interacting with adults: engaging actively with each other, requesting clarification, providing feedback, supporting each other's storylines and seeking to improve narrative coherence (Blum-Kulka, Huck-Taglicht, & Avni, 2004; Nicolopoulou & Richner, 2004; Schick & Melzi, 2010). These exchanges resemble the kinds of effective interactions described above, but are less readily facilitated in preschoolers' conversations with adults (Schick & Melzi, 2010). Importantly, such interactions can help children appropriate skills that mediate performance on individual non-verbal reasoning tasks (Rojas-Drummond, Pérez, Vélez, Gómez, & Mendoza, 2003; Wegerif, Mercer, & Dawes, 1999). There is, therefore, reason to believe that having preschoolers interact in small groups during a storytelling task without directive adult control could stimulate them to reason explicitly about their thinking and benefit from each other's knowledge and perspectives, which could in turn benefit their individual reasoning skills.

The present study

Bringing these perspectives together suggests that participating in shared narrative construction (i.e., storytelling) within a peer-group setting in the absence of directive adult control could stimulate causal reasoning in preschoolers. As far as we can determine, this combination of learning components has not previously been addressed in early childhood education research. It is therefore not known how preschoolers will talk with each other under these circumstances, nor whether this bears a relation to outcomes on causal reasoning tasks. The present study aims to answer these questions by investigating the kinds of discourse elicited when preschoolers take part in such a learning arrangement. Specifically, it examines whether different outcomes on causal reasoning tasks reflect differences in the types of causal reasoning and peer interaction exhibited during the storytelling task.

To this end, the *narrative* and *interactional* functions of children's utterances during shared storytelling were investigated. The approach taken was to compare discourse in the groups that showed the lowest and the highest improvement on causal reasoning tasks, after a storytelling intervention in multiple groups. This 'diverse-cases' approach compares cases (here: groups) representing the range of variation in outcomes (here: improvement in causal reasoning), in order to explore relationships between case characteristics (here: discourse functions) and the outcome variable (Gerring, 2007). The results could help generate hypotheses for future research on stimulating preschoolers' causal reasoning.

In this study, as in other studies investigating preschoolers' narrative abilities (e.g., Crawford & Hade, 2000; O'Neill et al., 2004), picture sequences from picture-books were used to create an open-ended process in which children assign meaning to visual stimuli. Picture-books are accessible, attractive and authentic materials, specially made for and familiar to young children. To date, picture-book research has largely focused on their potential to promote language and literacy functions and on parent-child interactions during shared reading (Mol, Bus, De Jong, & Smeets, 2008; Fletcher & Reese, 2005). Only a few studies have used picture-books to stimulate other cognitive and academic skills, such as promoting theory of mind understanding (e.g., Adrián, Clemente, & Villanueva, 2007; Symons, Peterson, Slaughter, Roche, & Doyle, 2005), understanding mathematical concepts (e.g., Anderson, Anderson, & Shapiro, 2005; Van den Heuvel-Panhuizen & Elia, 2011) and problem-solving (e.g., Gosen, 2012). Furthermore, most picture-book research focuses on interactions between child and adult (e.g., parent, teacher); only a few studies focus on interactions between children during picture-book activities. Thus, the present study also provides an original contribution to picture-book research in early childhood settings.

METHODS

Design

Three mainstream primary schools in the Netherlands with preschool classes took part in the study. Dutch preschool is generally organised as separate classes within primary schools, where children aged 4-6 learn through play-activities and other suitable experiences prior to the commencement of formal education.

The study employed a pretest/posttest group design and cross-case (group) comparison following the 'diverse-cases' approach described. Six storytelling groups (A to F) were formed (two per school), each consisting of 4-5 children. Within schools, groups were made as comparable as possible[1] on age, sex, pre-intervention estimators of general cognitive functioning and causal reasoning, home languages, parental level of education, home reading environment, and the extent of liking books and magazines (see Participants and Measurement sections). Descriptive statistics are in Table 2.1.

All groups participated in a two-week storytelling intervention. Pre-to-post intervention change on three causal reasoning tasks (see Measurement section) was aggregated across all individuals in a group. Kruskal-Wallis tests calculated group mean ranks for the aggregate change on each causal reasoning task; a total rank order for each group was then calculated (Table 2.2). The group with the lowest total rank order (i.e., Group A; 5 children) and the group with the highest total rank order (i.e., Group E; 4 children) were denoted as being the least-improving and the most-improving groups respectively. These groups were the subject of the cross-case comparison.

Participants

Participating schools received information packages for parents/caregivers of all children in their preschool classes, excluding children whom teachers knew to have behavioural and/or attentional deficits. The packages contained a letter about the

[1] Given the number of groups, group size and number of matching variables, significance testing of group differences on these variables is not meaningful.

Table 2.1
Group descriptive statistics pre-intervention

	School 1				School 2				School 3			
	Group A *N* = 5		Group B *N* = 5		Group C *N* = 5		Group D *N* = 5		Group E *N* = 4		Group F *N* = 4	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (months)	64.6	*7.1*	62.6	*6.5*	63.0	*4.3*	65.4	*5.1*	65.8	*4.7*	63.8	*3.3*
IQ estimate	79.6	*20.6*	72.2	*11.7*	112.8	*5.7*	115.6	*13.2*	99.3	*19.6*	101.3	*11.7*
Working memory	5.2	*3.3*	7.0	*3.5*	8.4	*1.8*	6.8	*1.1*	3.8	*2.5*	6.3	*1.9*
Attention	16.4	*1.7*	12.4	*14.2*	19.6	*0.9*	19.2	*1.8*	17.0	*4.7*	19.8	*0.5*
Home reading environment	4.8	*1.5*	4.8	*2.3*	7.4	*2.5*	8.2	*1.8*	5.5	*2.1*	6.0	*2.2*
Liking for books & mags	2.8	*0.4*	2.6	*0.5*	2.2	*0.8*	3.0	*0.0*	2.8	*0.5*	2.5	*0.6*
Physical causality	1.2	*1.6*	1.8	*1.8*	1.4	*1.1*	2.0	*1.4*	0.8	*1.0*	1.3	*1.5*
Psychological causality	2.6	*0.5*	3.0	*0.0*	2.6	*0.5*	3.0	*0.0*	2.0	*0.8*	2.5	*0.6*
Situational causality	4.2	*0.8*	4.0	*0.7*	4.8	*0.4*	3.8	*0.4*	4.0	*0.8*	4.3	*1.0*

Table 2.2
Pretest-to-posttest change on causal reasoning tasks

		Physical causality		Psychological causality		Situational causality		Kruskal-Wallis
Group	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	Total Rank
A	5	0.20	*1.30*	0.00	*0.71*	-0.20	*1.30*	37.0
B	5	0.60	*1.14*	-0.40	*0.89*	0.60	*0.89*	41.8
C	5	0.60	*1.14*	0.20	*0.45*	-0.40	*0.89*	39.1
D	5	1.20	*0.45*	0.00	*0.00*	0.20	*0.45*	45.6
E	4	1.75	*1.50*	0.75	*0.96*	0.25	*0.50*	54.4
F	4	1.25	*1.26*	0.25	*0.50*	0.00	*0.82*	45.8

purpose of the study, a request to participate, an informed consent form, and a questionnaire containing demographic items (e.g., child's sex and age, home languages, parental level of education). The questionnaire also requested information about the home reading environment and the child's liking for books and magazines. Home reading environment was operationalised as the extent to which different types of books and magazines (i.e., story books, first picture-books, picture-books for preschoolers, early readers, chapter books, information books, comic books, children's magazines, other children's books) were available in the home. Liking for books and magazines was indicated on a 3-point scale as 'does not like', 'likes a bit' or 'likes a lot'.

Twenty-eight children who were comparable with other children at their school on age, background variables and pretest scores were selected to participate in the storytelling intervention. Of these, $N = 15$ (54%) were boys. Sample mean age was 64 months ($SD = 5$m; range 54-72m). Within each school, children were assigned to one of two storytelling groups.

Materials

Two picture-books were used: the Dutch versions of *Cave baby* (Donaldson & Gravett, 2010) and *Tom's elephant* (Harvey & Klaassen, 1997) (see Appendix 2.A). Books were selected on five criteria. First, the story and illustrations were judged in children's literature reviews to be of high quality, captivating and attractive to both young children and adults. Second, a clear storyline could be derived from the illustrations. Parts of the storyline were amenable to different interpretations, however, which is considered to be a requirement of group tasks at primary school level (Howe & Mercer, 2010). Third, the sequence of illustrations depicted situations that could be explained by constructing a causal chain of events (i.e., a given situation could be understood from information depicted in earlier situations). Fourth, the stories were fictional but situated in meaningful contexts governed by natural laws. This precluded the use of books where magical rather than causal forces could be used to explain the chain of events. Fifth, the selected books were novel to all participants (as verified by teachers), to ensure that earlier experiences did not affect the results.

Procedure

Groups were supervised by three research assistants who were final-year Master students in Pedagogical Sciences and fully trained in using the materials and protocols. Each group completed three sessions per week for two weeks in a quiet room at school. The research assistants encouraged all children to participate but - importantly - did not intervene with respect to what was said. A detailed protocol covered what instructions should be given, how to display the illustrations, how long to wait before turning a page, how to support participation, etcetera.

In the first two sessions (week 1), the children were asked to generate a storyline together while looking at the illustrations of *Cave baby*. No child was able to read the text of the book. The third session included a hand-puppet as a prop. Following O'Neill et al. (2004), the children were told that the puppet loved stories and had not heard this one before. They were then asked to tell it the story while looking at the illustrations. The presence of a 'naïve' listener can be an extra stimulus for children to articulate their reasoning more than they would to a knowledgeable listener. This procedure was repeated in week 2 using *Tom's elephant*.

Importantly, children were allowed as much time as they wished to tell the stories, so that each group ultimately spent a different amount of time on the intervention. The first two sessions lasted 25 minutes on average (range 13-29); the third session lasted 22 minutes on average (range 7-34). Storytelling sessions were digitally recorded. Six recordings (17%) were lost due to technical failures.

Measurement

Pretests and posttests

Children were individually tested on general cognitive functioning and causal reasoning (i.e., *physical*, *psychological* and *situational* causality). Pretests were performed an average of 23 days (range 11-34) before the intervention and posttests an average of 11 days (range 3-33) after. Children were tested by a research assistant in a quiet room. Test sessions lasted 49 minutes on average (range 30-65).

General cognitive functioning was measured using well-known, validated tests, administered following their standard protocols. To estimate intellectual abilities, the Vocabulary and Block Design subtests of the Dutch version of the Wechsler Preschool and Primary Scale of Intelligence - Third Edition (WPPSI-III-NL; Hendriksen & Hurks, 2009) were administered. Combining these subtests gives a reliable estimate of children's intelligence (Sattler & Dumont, 2004). The Vocabulary subtest comprises 25 items to be named or defined (max raw score 45). The Block Design subtest requires 20 visually presented patterns to be reproduced (max raw score 40). Raw scores for the two subtests were transformed to standard scores and a total IQ estimator calculated following the standard procedure. Working memory was assessed with a computerised visual-spatial grid task in which sequences of dots have to be remembered and reproduced (Klingberg, Forssberg, & Westerberg, 2002; max score 16). Attention was measured as accuracy on the 'Cats' task of the NEPSY visual attention subtest, which requires the subject to search for a target embedded in a random array of different pictures within a set time (Korkman, Kirk, & Kemp, 1998; max raw score 20).

Physical causality was measured with an instrument derived from Das Gupta and Bryant (1989). Pairs of picture sequences were presented, each showing an object in an initial state followed by that object in an end state. The initial state was non-canonical in one way (e.g., a wet cup or a broken cup), while the end state was non-canonical in two ways (e.g., a wet, broken cup). The child had to select one of two agents (e.g., faucet or hammer) that could cause the end state from the initial state. Both agents are needed to cause the end state from the canonical state of the object; however only one is needed to cause the end state from an initial state in which one transformation has already taken place. In order to be correct on both sequences in a pair, children needed to make a causal inference about the change between the initial and end state of the depicted object. In the example, the hammer is the agent for the sequence *wet cup - wet, broken cup* and the faucet is the agent for the sequence *broken cup - wet, broken cup*. In total, eight sequences were presented (i.e., four pairs). The two sequences of a pair were never presented consecutively. The number of pairs for which both sequences were correct was summed to give a total score (max 4).

Psychological causality was measured with an instrument developed by the research team. Children were presented with a picture of a psychological state and had to indicate which of three other pictures showed a causal antecedent. For example, an angry rabbit was shown. Possible causal antecedents showed the rabbit raking a pile of leaves together in the presence of another character who was acting in three different ways: (a) lying calmly; (b) jumping onto the pile of leaves; (c) moving out of the scene. The instrument comprised three such items.

Situational causality was measured with a similar 5-item instrument developed by the research team. Children were presented with a picture of a situation and had to indicate which of three other pictures showed a causal antecedent. For example, a character was shown covered in cake-mix and throwing back his arms in surprise with an upturned mixing bowl on a table in front of him. Possible causal antecedents were: (a) a ball heading straight for the bowl; (b) a stationary ball under the table; (c) a ball knocking the table over and the contents of the bowl to the ground.

Group discourse

Discourse in the groups showing the lowest and the highest improvement in causal reasoning post-intervention (i.e., Groups A and E respectively) was investigated through qualitative content analysis, involving systematic observation, coding and analysis of children's verbal utterances in terms of their narrative and interactional functions. Seven mutually exclusive *narrative* categories (i.e., labelling, observation, elaboration, inference, prediction, explanation, other) captured types of causal reasoning as well as types of non-causal narrative utterances found in previous research (e.g., Hickling & Wellman, 2001; Walsh, 2003). Definitions and examples of each category are given in Appendix 2.B.1 (note that counterfactual thinking was not observed).

Interactional categories were developed to reflect types of interaction between children. An initial set of categories was derived from previous research on children's talk in primary school classrooms and during causal reasoning tasks specifically (Howe & Mercer, 2010; Teasley, 1997). The initial categories were used to code the longest transcript for each group. It became clear that not all categories were relevant for the present data and that distinctions between some categories were too fine to be applied reliably. Consequently, the set was reduced to seven mutually exclusive categories (i.e., assertion, extension, cumulative, disputational, clarification, challenge, other). Definitions and examples of each category are given in Appendix 2.B.2.

All video-recordings for both groups were transcribed and coded. One Group A recording was missing. Transcriptions excluded utterances unrelated to the task. Utterances were segmented into speech units, defined as a single meaningful unit of information serving one or more functions of interest to the study. Each speech unit was assigned a single narrative category and a single interactional category; thus each speech unit had exactly two functional codes. Coding reliability was confirmed for 25% of the transcribed corpus by a colleague who was trained in using the coding protocols and who was blind to the study hypotheses and the group differences. Percentage agreement for both types of codes was 80%. Inter-rater reliability was further tested using Krippendorff's α-coefficient for nominal data; coefficients of .72 for narrative codes and .76 for interactional codes indicate acceptable reliability.

Cross-case analysis

Discourse in the groups showing the lowest and the highest improvement in causal reasoning post-intervention (i.e., Groups A and E respectively) was analysed. Discourse duration was calculated as the total amount of time spent on the intervention across all six sessions. Discourse participation of individual group members was calculated as a proportion of the total number of speech units per group. Comparison of group discourse was done by aggregating the number of occurrences of each coding category across sessions per group and obtaining patterns of narrative and interactional talk from the proportions represented by each category. These patterns were compared using the Chi-square test and Fisher's exact tests compared individual categories between the two groups. Analyses were performed in IBM SPSS® Statistics 20 ($\alpha = .05$).

RESULTS

Discourse duration and participation

The groups spent notably different amounts of time on the intervention: Group A (i.e., lowest improvement) spent a total of 1 hour 23 minutes (M = 14 min/session) while Group E (i.e., highest improvement) spent 2 hours 38 minutes (M = 26 min/session).

In Group A, the five members were respectively responsible for 29.6%, 26.2%, 25.7%, 16.6% and 1.7% of speech units. In Group E, the four members were respectively responsible for 37.3%, 30.1%, 18.9% and 13.7% of speech units.

Narrative talk

Group A produced mainly *labelling* (31.7%), *inferences* (24.2%) and *other* statements (21.8%). Group E exhibited talk that was more diverse and balanced, with substantial proportions of *labelling* (17.8%), *observations* (15.0%), *inferences* (17.5%) and *other* statements (26.6%). Group patterns were significantly different ($\chi^2(6) = 98.78$, $p < .001$), with a higher proportion of *labelling* and *inferences* in Group A (both $p < .001$), and a higher proportion of *observations* ($p < .001$), *explanations* ($p < .001$) and *other* statements ($p = .02$) in Group E. Group differences were not significant for *elaborations* and *predictions*. Results summarised across all sessions are presented in Figure 2.1 and session statistics are reported in Appendix 2.C.1.

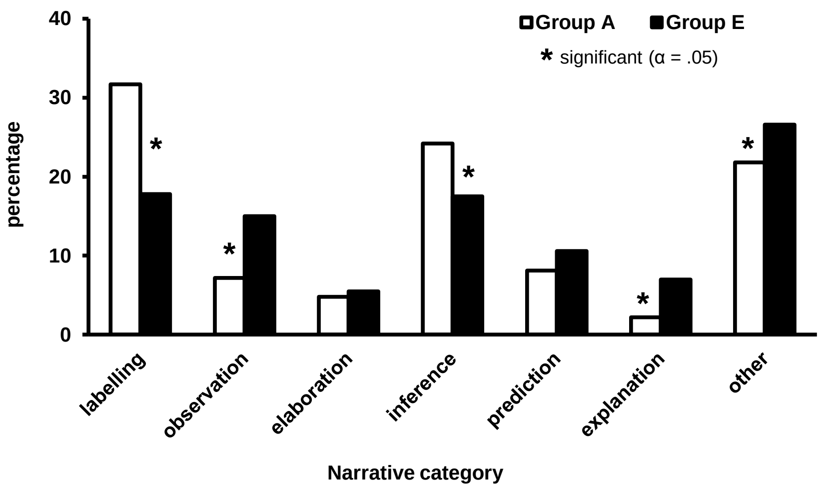

Figure 2.1. Narrative category percentage occurrence across sessions

Excerpts from the transcriptions are illustrative of these differences. First, much Group A discourse went no further than identifying perceptual features of the stimulus (i.e., *labelling*). By comparison, Group E discourse contained more detailed descriptions or interpretations of perceptual features (i.e., *observations*). The following example from *Cave baby* concerns stimuli representing a sabre-toothed tiger and a hyena. Group A statements concerning these stimuli were typically: “*I see a tiger*”, “*A tiger and an elephant*”, “*This is a tiger with sharp teeth*”, “*A zebra*” (i.e., the hyena), “*That’s not a zebra, that’s a wolf*”, “*That’s a long tiger*”. Group E, on the other hand, made several observations focusing on the tiger's teeth: "*There's a great big tiger with a sharp tooth*", "*He is big and the tooth is big but this is a bit smaller and his body is bigger than the tooth*", "*Do you know how big the teeth are? Longer than...*", "*But this is longer than this and big, but this is even longer*". Then, when they encountered the picture of the hyena and took it to represent the same tiger, the children continued to focus on its teeth: "*But his big teeth are a bit broken now*". This then led them to offer an explanation of why the teeth were smaller in the picture of the hyena, during which they invoked the causal mechanism that chewing grass breaks teeth. Thus, producing more detailed descriptions and interpretations of perceptual features may stimulate causal reasoning by drawing attention to situations or events that invite or need explanation.

A second difference concerns the use of *inferences* (i.e., connections between situations or events inferred from the stimulus) and *explanations* (i.e., references to causal mechanisms). In *Tom's elephant*, the stimulus is a picture of the elephant getting free of shackles after being chained up. Group A discourse consisted of several simple inferences:

Child 1: *The elephant is free.*
Child 2: *He broke everything.*
Child 2: *He sucked it all up with his trunk.*
Child 3: *Yes, and by stamping.*

In Group E, discourse about the same stimulus went as follows:

Child 1: *Maybe he wants a strong wind, maybe he stamped really hard.*
Child 2: *Look here, do you know why this opened? Look, he stamps really hard and then he goes there to that side and then he also goes again.*
Child 3: *I think that he sees water there and that he...*
Child 3: *And that he, no, that his legs are really strong so the cuffs around his leg break.*
Child 4: *Look, the elephant has strong claws and then he goes like boom boom and breaks the locks and that iron.*

These excerpts show that, while Group A's causal reasoning was limited to drawing simple inferences about the stimuli, Group E engaged in reasoning about the causal mechanisms that could underlie the depicted situation.

Interactional talk

Almost a third of all statements in both groups was in the form of self-contained *assertions*. Both groups also produced a substantial proportion of *extensions* (Group A: 16.7%; Group E: 19.7%). Group A was further characterised by *cumulative* (15.5%) and *disputational* (12.1%) interactions, while Group E produced a sizeable proportion of *clarifications* (13.9%). Group patterns were significantly different ($\chi^2(6) = 58.16$,

$p < .001$). There was a higher proportion of *cumulative* and *disputational* talk in Group A ($p < .001$ and $p = .003$ respectively), while Group E produced more *clarifications* ($p < .001$). Group differences were not significant for *assertions*, *extensions*, *challenges* or *other* interactions. Results summarised across all sessions are presented in Figure 2.2 and session statistics are reported in Appendix 2.C.2.

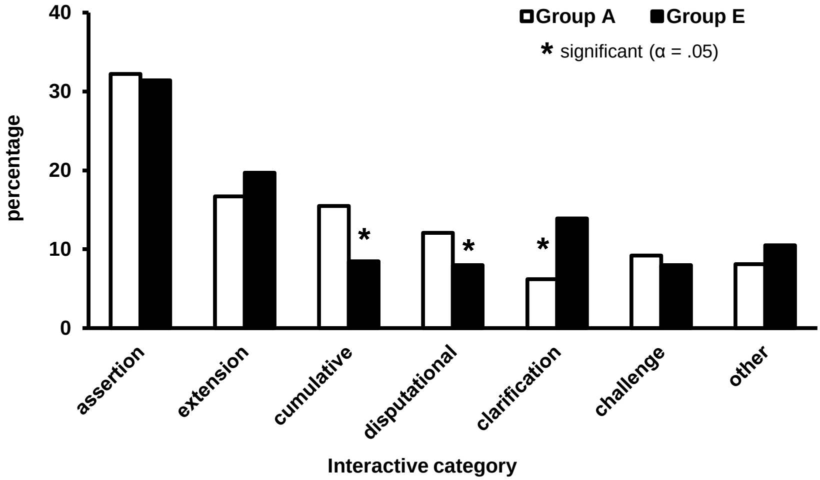

Figure 2.2. Interactional category percentage occurrence across sessions

To illustrate: in *Cave baby*, the stimulus was a picture of the baby and the mammoths playing amidst big splashes of colour. Group A discourse was as follows:

Child 1: *They're setting off fireworks.*
Child 2: *No.*
Child 1: *Yes!*
Child 3: *No, they are painting.*
Child 1: *No, fireworks!*
Child 4: *Painting!*
Child 1: *Fireworks!*
Child 4: *Painting!*
Child 1: *Fireworks, fireworks, fireworks!*
Child 2: *Yes, they are painting.*
Child 3: *Yes.*
Child 1: *No!*
Child 3: *Yes!*
Child 1: *They are painting and setting off fireworks.*
Child 1: *They're setting off painted fireworks.*

Child 3: *Yes, that could be.*

Child 1: *I'm right!*

Clearly, little or no causal reasoning is stimulated by these interactions, which contain little more than confirmations and repetitions (i.e., *cumulative* talk), simple rebuttals and reassertions (i.e., *disputational* talk). By comparison, an excerpt illustrating the *clarifications* distinguishing Group E involves a discussion about whether the mammoth is using his trunk to point with in one scene of *Cave baby*:

Child 1: *He wants to go to the baby otherwise he'll run away.*

Child 2: *Look, look, this is pointing there.*

Child 1: *Yes, because he's got hands but then he'll fall on his behind.*

Child 2: *No.*

Child 1: *Yes, he's got hands but they're a bit flat here* [gestures] *and we've got hands like this* [gestures].

Child 2: *I don't know what she means.*

Child 3: *Really?*

Child 3: *Look, he's pointing with his trunk.*

Child 3: *Look, look, then he can't really do like this...so.*

Child 4: *I think it's pointing to him.*

This excerpt shows Group E children explicitly justifying their statements. Starting from the notion that the trunk could be pointing, Child 1 tries to justify (anatomically) why the mammoth would point with his trunk rather than with his feet (here, hands), and is stimulated to become more explicit by Child 2 not accepting her arguments. As seen here, clarifications may involve causal reasoning being made visible and may result in agreement being reached.

DISCUSSION

This study describes discourse elicited when preschoolers participate in undirected, shared picture-book storytelling within a peer-group setting. Based on previous research, the learning arrangement was expected to stimulate children to form coherent, causally-sequenced plots whilst sharing their knowledge, experience and perspectives, which could help to improve causal reasoning skills. The study examined whether different outcomes on causal reasoning tasks reflect differences in the types of causal reasoning and peer interaction exhibited during the storytelling task.

Twenty-eight children from three preschools were placed in six groups and asked to generate narratives together from picture-book stimuli in the absence of directive adult control of the discourse. Discourse was then compared in the two groups that showed the lowest and the highest improvement on causal reasoning tasks post-intervention. The main differences between these groups are presented in Box 1.

Box 1: Group discourse differences

Narrative statements:	Least-improving group produced more labelling and simple inferences Most-improving group produced more detailed observations and causal explanations
Interactional statements:	Least-improving group produced more disputational and cumulative exchanges Most-improving group offered or requested information justifying or clarifying previous statements
Discourse duration:	Most-improving group spent twice as long on the intervention

Regarding narrative statements, relatively more simple inferences were produced by the least-improving group, while relatively more causal explanations were produced by the most-improving group, often in relation to detailed interpretations of perceptual features. There were no group differences in causal predictions. This suggests that engaging in explanations invoking causal mechanisms could play a key role in improving causal reasoning. This lends support to previous findings regarding the importance of children's explanations in learning and thinking. Causal explanations - particularly of surprising events or inconsistent outcomes - are thought to be especially powerful in influencing children's causal reasoning and causal learning (Legare et al., 2010; Wellman & Liu, 2007), and young children make use of causal-explanatory theories to make sense of real-life happenings (Hickling & Wellman, 2001).

Intriguingly, the groups displayed patterns of interaction that show considerable similarities to patterns found with older children (e.g., Howe & Mercer, 2010; Rojas-Drummond et al., 2003; Wegerif et al., 1999). Children in the most-improving group engaged critically but constructively with each other's ideas and explicitly offered justifications of their statements for joint consideration. The least-improving group was distinguished by exchanges typifying what is known as disputational and cumulative talk, both of which are associated with lower learning outcomes (Howe & Mercer, 2010). Disputational talk is characterised by disagreement and exchanges of assertions and counter-assertions. With cumulative talk, speakers add uncritically to another's statements by repeating, confirming or elaborating on what has gone before. Taken together, these findings suggest that group discourse beneficial to causal reasoning in preschoolers may be characterised by the kinds of interactions known as 'exploratory talk', but also support the view that placing children in groups does not guarantee exchanges that are likely to enhance learning (Howe & Mercer, 2010).

The design allowed differing amounts of time to be spent on the intervention, as determined by the children themselves. Notably, the most-improving group spent the most time on the intervention, while the least-improving group spent the least. This strongly suggests that duration of shared storytelling experiences could be influential. Specifically, more productive types of narrative thinking or peer interaction could emerge only in later stages of discourse.

Finally, it is worth noting that discourse participation was not balanced: discourse was dominated by two or three group members while others contributed much less. This concurs with previous findings on unstructured discussions (Howe & Mercer, 2010). Although a lack of active contribution does not preclude children from benefiting from the knowledge and perspectives made visible by others (e.g., vicarious learning, observational learning), the relationship between active participation and learning in the present study is unclear.

Research implications

As stated, this may be the first study investigating how undirected shared picture-book storytelling in a peer-group setting could stimulate causal reasoning in preschoolers. While its small-scale nature does not permit conclusions to be drawn about the effects of the learning arrangement, it does allow detailed, in-depth examination of the discourse that takes place. This, in turn, helps generate hypotheses that should be tested in future research on a larger scale and using a broader range of causal reasoning tasks.

First, it appears that discourse involving causal explanations and detailed interpretations of perceptual features is beneficial to causal reasoning. The question is, therefore, how to stimulate children to move past superficial reactions to deeper treatment of the material. Second, it appears that preschoolers can interact in ways similar to exploratory talk taught to older children and that this benefits their learning. It would therefore be of interest to investigate the extent to which preschoolers can explicitly learn and use ground-rules for exploratory talk and the effects on learning. Third, more time spent on storytelling appears to be related to better outcomes, possibly because discourse becomes more productive after some time has elapsed. Research focusing on the temporal dimension of discourse progression with preschoolers should investigate this issue (Mercer, 2008).

If this hypothesis is confirmed, it would be important to investigate how children could be engaged in the intervention for longer periods of time. The 'travelling lens' model of attention used in literacy and television-viewing research (e.g., Linebarger, Moses, Garrity Liebeskind, & McMenamin, 2013) posits that children's interest in and attention to stimuli depends on stimulus features (e.g., novelty, complexity, recognisability) in relation to individual dispositions. This gives pointers for investigating how many sessions - as a function of picture-book characteristics and individual or group dispositions - are optimal for arousing and sustaining children's interest and attention and for manipulating the complexity and novelty of stimuli. For example, picture-books could be used that present multiple storylines simultaneously, so that children have to move backward and forward through the book to seek connections between story elements (Wolfenbarger & Sipe, 2007).

The study did not examine the influence of individual differences on the results. However, important differences between children - related to individual characteristics, social histories and past experiences - could affect how they participate in shared storytelling and the discourse patterns shown (Howe & McWilliam, 2001; Howe & Mercer, 2010; Schick & Melzi, 2010; Wellman & Liu, 2007). For example, children's prior knowledge of and experience with books and storytelling affects how they make sense of picture-book stimuli (Crawford & Hade, 2000), while the extent to which their social experiences stimulate reasoned argumentation affects how they negotiate and justify their reasoning in interactions with peers (Ehrlich, 2011; Howe & McWilliam, 2001; Howe & Mercer, 2010). It is important to untangle these influences in future studies.

Finally, future research should examine the effects of increasing the adult's role in the discourse, for example through the kinds of strategies used in research with older children (Mercer & Howe, 2012). It is likely that adult intervention is needed to help groups engage in types of discourse that are beneficial to learning outcomes (e.g., detailed observations, causal explanations, justification) and avoid those that are not (e.g., labelling, simple inferences, repetition, disagreement). Specifically, adults may need to guide children towards potentially relevant cues, for example by encouraging them to explain story events (Girard, Girolametto, Weitzman, & Greenberg, 2013) and by stimulating the acquisition of effective discourse skills.

Educational implications

The present study contributes to a line of research that aims to provide early childhood educators with accessible and low-threshold methods and materials that stimulate young children's thinking skills (e.g., Bodrova & Leong, 2007; Diamond, Barnett, Thomas, & Munroe, 2007). This could help to improve young children's school readiness, which is an issue of major importance to educators, public and politicians (e.g., National School Readiness Indicators Initiative and Head Start program in the US, Head Start in Australia and Sure Start programme in the UK).

In this regard, it is important to note that fully authentic materials - familiar and accessible to both early childhood educators and young children - were used here.

Picture-books were selected on the basis of criteria that non-researchers could easily learn to apply and were not shortened or adapted, as is often the case in educational research. Moreover, the small group setting increases the ecological validity of the approach, as schools rarely have sufficient resources to carry out such activities on a one-to-one basis.

Finally, the role of the adult in this study was to create conditions for active, open and equitable group discourse rather than to direct or control it, an approach shown to facilitate learning in peer-groups (Mercer & Howe, 2012). If future research confirms that these conditions are effective for preschoolers, this could have pedagogical implications. First, early childhood educators need to include skills for facilitating this type of discourse in their repertoire. Second, it may be difficult for educators with more directive styles to resist the temptation to control children's discourse in their habitual ways (Howe & Mercer, 2010). In that case, they need to learn how to manage discourse that may need guidance in form but freedom in content.

Acknowledgements

This research was partially supported by a grant from the Dutch National Initiative Brain and Cognition (NIBC; project number 056-31-013). The authors thank especially Marijn Knibbeler and Marije Broens-Paffen, Anneleen Post, Shanna Anema, Evelien Broersen, Femke Kramer, Anton Boonen and the participating schools, caregivers and children for their contributions.

Appendix 2.A. Picture-book materials

'Cave Baby' (Donaldson & Gravett, 2010) is situated in prehistoric times, where a baby lives with his parents in a cave. His mother paints animals on the walls of the cave. Cave Baby decides to join in by adding splashes of paint to the walls; this angers his parents, however. Come night, Cave Baby is kidnapped by a mammoth who takes him on a ride through a moonlit landscape populated by different animals to his own cave. There, Cave Baby happily paints and plays with the mammoths. Finally, the mammoth takes Cave Baby back home to bed. The book comprises 32 pages with colour illustrations.

In 'Tom's Elephant' (Harvey & Klaassen, 1997), a circus comes to Tom's town. Tom is awakened one night by a strange noise. An elephant has run away from the circus and is hiding in the garage. Tom and the elephant become friends and the elephant helps Tom to do his paper rounds and punishes the school bully by putting him up in a tree. Tom tries to hide the elephant but the police find him and lock him up in chains. A violent storm arises and the town is flooded. The elephant escapes and rescues people from the water. The townsfolk hail him as a hero and send him on a boat back home to Africa. The book comprises 25 pages with colour illustrations.

Appendix 2.B. Example narrative and interactional categories

Table 2.B.1
Narrative categories

Category	Definition	Examples
Labelling	Naming or identifying perceived features of the stimulus without further information	"*There is paint*"; "*The baby is waving*"
Observation	Description or interpretation of perceptual features that goes beyond labelling	"*He is big and the tooth is big but this is a bit smaller and his body is bigger than the tooth*"
Elaboration	Information about the stimulus that represents prior knowledge	"*Look, you can fold up these steps and then you can drive it* [i.e., a caravan] *with a car*"; "*Zebras don't have hair there*"
Inference	Statement about a connection between actions, events, characters or psychological variables that is inferred from the stimulus	"*They want to kill the mouse*"; "*The baby's frightened of the elephant*"
Prediction	Expression of an expected outcome of a depicted or inferred action or event	"*The elephant's going to pull him out of there*"; "*The mother is going to think that the baby made those tracks*"
Explanation	Reference to a causal mechanism to explain a depicted or inferred action, event or psychological state	"*Do you know why the window is open? Because the elephant stamps really hard and makes a wind that blows there*"; "*They're walking next to the father so that's why you can see them in the water*"
Other	Non-specific, non-narrative or incomprehensible statement	"*Wow!*"; "*I think the baby is cute*"; "*What do you mean?*"; "*Yes*"

Note. Counterfactuals were not observed and are therefore not included.

Table 2.B.2

Interactional categories

Category	Definition	Examples
Assertion	Self-contained claim about a story element	"*That's a circus for sure*"
Extension	Additional thought about or elaboration on a previous statement	Second part of the exchange: "*He is looking behind him*"/"*He doesn't want to be seen and has to hide somewhere*"
Cumulative	Statement that builds positively but uncritically on an earlier statement, characterised by confirmation, repetition or completion of an unfinished statement	"*He is right*"; second part of the exchange: "*And then he has a fright so he lets go of his torch*"/"*Then he has a fright and drops his torch*"; second part of the exchange: "*That's a big....*"/"*Bath*"
Disputational	Disagreement in the form of a simple rebuttal or reassertion without further justification or elaboration	"*No it isn't*"; "*Yes it is*"
Clarification	Provides or requests information justifying or clarifying a previous statement	"*It's a circus because there's a clown here*"; "*What do you mean?*"
Challenge	Alternative interpretation of a previous statement accompanied by a justification of that interpretation	"*No, it's a swimming pool* [as opposed to a ditch] *because there's a pool-ladder there*"
Other	Non-specific or incomprehensible statement	"*Woh!*"

Appendix 2.C. Session statistics

Table 2.C.1

Frequency and proportion of narrative categories per session

Category	Session 1	Session 2	Session 3	Session 4	Session 5[a]	Session 6	Total
			Group A				
Labelling	71 (48.0%)	31 (18.0%)	34 (33.0%)	53 (30.3%)	-	56 (31.8%)	245 (31.7%)
Observation	7 (4.7%)	15 (8.7%)	8 (7.8%)	18 (10.3%)	-	8 (4.5%)	56 (7.2%)
Elaboration	9 (6.1%)	6 (3.5%)	1 (1.0%)	9 (5.1%)	-	12 (6.8%)	37 (4.8%)
Inference	28 (18.9%)	38 (22.1%)	27 (26.2%)	52 (29.7%)	-	42 (23.9%)	187 (24.2%)
Prediction	7 (4.7%)	14 (8.1%)	9 (8.7%)	11 (6.3%)	-	22 (12.5%)	63 (8.1%)
Explanation	2 (1.4%)	4 (2.3%)	2 (1.9%)	4 (2.3%)	-	5 (2.8%)	17 (2.2%)
Other	24 (16.2%)	64 (37.2%)	22 (21.4%)	28 (16.0%)	-	31 (17.6%)	169 (21.8%)
Total	148	172	103	175	-	176	774
			Group E				
Labelling	66 (20.9%)	52 (20.1%)	15 (11.4%)	25 (18.2%)	28 (17.4%)	10 (10.5%)	196 (17.8%)
Observation	49 (15.5%)	47 (18.1%)	25 (18.9%)	15 (10.9%)	14 (8.7%)	15 (15.8%)	165 (15.0%)
Elaboration	24 (7.6%)	8 (3.1%)	6 (4.5%)	6 (4.4%)	13 (8.1%)	3 (3.2%)	60 (5.5%)
Inference	46 (14.6%)	38 (14.7%)	21 (15.9%)	25 (18.2%)	33 (20.5%)	29 (30.5%)	192 (17.5%)
Prediction	16 (5.1%)	36 (13.9%)	26 (19.7%)	11 (8.0%)	18 (11.2%)	10 (10.5%)	117 (10.6%)
Explanation	15 (4.7%)	19 (7.3%)	11 (8.3%)	8 (5.8%)	19 (11.8%)	5 (5.3%)	77 (7.0%)
Other	100 (31.6%)	59 (22.8%)	28 (21.2%)	47 (34.3%)	36 (22.4%)	23 (24.2%)	293 (26.6%)
Total	316	259	132	137	161	95	1100

Note. [a] Group A recording lost.

Table 2.C.2

Frequency and proportion of interactional categories per session

Category	Session 1	Session 2	Session 3	Session 4	Session 5[a]	Session 6	Total
			Group A				
Assertion	49 (33.1%)	50 (29.1%)	32 (31.1%)	55 (31.4%)	-	63 (35.8%)	249 (32.2%)
Extension	19 (12.8%)	20 (11.6%)	22 (21.4%)	38 (21.7%)	-	30 (17.0%)	129 (16.7%)
Cumulative	27 (18.2%)	25 (14.5%)	14 (13.6%)	24 (13.7%)	-	30 (17.0%)	120 (15.5%)
Disputational	15 (10.1%)	32 (18.6%)	14 (13.6%)	17 (9.7%)	-	16 (9.1%)	94 (12.1%)
Clarification	13 (8.8%)	10 (5.8%)	7 (6.8%)	11 (6.3%)	-	7 (4.0%)	48 (6.2%)
Challenge	13 (8.8%)	15 (8.7%)	7 (6.8%)	20 (11.4%)	-	16 (9.1%)	71 (9.2%)
Other	12 (8.1%)	20 (11.6%)	7 (6.8%)	10 (5.7%)	-	14 (8.0%)	63 (8.1%)
Total	148	172	103	175	-	176	774
			Group E				
Assertion	65 (20.6%)	80 (30.9%)	56 (42.4%)	47 (34.3%)	61 (37.9%)	36 (37.9%)	345 (31.4%)
Extension	60 (19.0%)	41 (15.8%)	33 (25.0%)	27 (19.7%)	30 (18.6%)	26 (27.4%)	217 (19.7%)
Cumulative	34 (10.8%)	24 (9.3%)	10 (7.6%)	6 (4.4%)	14 (8.7%)	6 (6.3%)	94 (8.5%)
Disputational	33 (10.4%)	25 (9.7%)	4 (3.0%)	9 (6.6%)	9 (5.6%)	8 (8.4%)	88 (8.0%)
Clarification	56 (17.7%)	41 (15.8%)	11 (8.3%)	21 (15.3%)	20 (12.4%)	4 (4.2%)	153 (13.9%)
Challenge	33 (10.4%)	24 (9.3%)	9 (6.8%)	5 (3.6%)	12 (7.5%)	5 (5.3%)	88 (8.0%)
Other	35 (11.1%)	24 (9.3%)	9 (6.8%)	22 (16.1%)	15 (9.3%)	10 (10.5%)	115 (10.5%)
Total	316	259	132	137	161	95	1100

Note. [a] Group A recording lost.