



Department of Psychology

**The Effects of Peer Discussion and Question Format on Realism in
Confidence Regarding Children's Testimonies in a Co-witness Situation**

Sandra Buratti and Sandra MacLeod

Master Thesis

Supervisor:

Carl Martin Allwood

Abstract

This experimental study examined the effect of peer discussion and the effect of question format: free recall, follow-up and focused questions, on children's (age of 9-11 years) memory recall of a witnessed event. The children were randomly assigned in pairs that watched one of two perspectives of a film. Half of the pairs discussed the film and the other half of the pairs had a neutral discussion. The results indicated that the children were less accurate, less confident and showed poorer realism in their confidence during follow up questions compared to free recall. Girls continued to show perfect realism during follow up. No effect of peer discussion was found on any of the dependent measures. However, 28 % of the children in the film discussion condition reported events that they could not have seen. The children were just as confident about these statements as they were about other statements they made during the interview.

Keywords: Eyewitness memory, Event recall, Co-witness, Confidence, Discussion, Children

Acknowledgements

We would like to thank all the children, parents, teachers and principals at the five different schools that participated in this study. Without your willingness and positive attitude this study would not have been possible.

We would also like to thank Professor Carl-Martin Allwood for inviting us to write this paper. Thanks to your guidance, insightful suggestions and patience we have been able to produce a thesis that give us pride.

We would like to thank Detective Inspector Karin Geistrand at Simrishamn Police, for her valuable insights in the realities of forensic interviewing. We are also grateful to all our fellow students who took time from their own thesis to help us produce the experiment film. We would also like to thank the children that helped us to develop the testing instructions and questionnaires.

I, Sandra B, would also like to thank my “partner in crime”, Sandra M, for bringing humor and patience to our collaboration. I can’t believe we made it! I would also like to thank my parents and my sister for their unconditional love and support. Without you I would be lost beyond words.

I, Sandra M would like to thank Sandra B, for her devoted work and friendship and for getting everything done in time.

Table of Contents

| | |
|---------------------------------------------------------------------------------------|----|
| Abstract | 2 |
| Acknowledgements | 3 |
| Background | 6 |
| <i>Children's Event Recall Memory</i> | 7 |
| <i>Different Question Formats</i> | 8 |
| <i>Metamemory, Confidence and Realism in Confidence</i> | 10 |
| <i>Confidence Judgments and Gender Differences</i> | 12 |
| <i>The Suggestibility of Children</i> | 12 |
| <i>Misinformation and Confidence in Children</i> | 15 |
| <i>Peer Discussion and the Incorporation of Post-event Information</i> | 16 |
| Hypotheses | 19 |
| Method | 20 |
| <i>Participants</i> | 20 |
| <i>Design</i> | 21 |
| <i>Material</i> | 22 |
| <i>Video Clip</i> | 22 |
| <i>Equipment</i> | 22 |
| <i>Questions used in the Discussion Conditions</i> | 23 |
| <i>Questionnaire for the Focused Questions and Confidence Scales</i> | 23 |
| <i>Preparation of the Free Recall and Follow Up Material for Confidence Judgments</i> | 24 |
| <i>Procedure</i> | 25 |
| <i>Session One</i> | 25 |
| <i>Session Two</i> | 26 |
| <i>Scoring Procedure and Coding</i> | 28 |
| <i>Calibration Measures</i> | 28 |

| | |
|--------------------------------------------------------------------|----|
| Results | 29 |
| <i>Analyzes of Peer Discussion and Questions Formats</i> | 29 |
| <i>Preliminary Analysis</i> | 29 |
| <i>General Analysis</i> | 31 |
| <i>Completeness.</i> | 31 |
| <i>Accuracy.</i> | 31 |
| <i>Confidence.</i> | 31 |
| <i>Calibration curves.</i> | 32 |
| <i>Bias.</i> | 34 |
| <i>Slope.</i> | 35 |
| <i>Analyzes of Critical Events and Question Format</i> | 35 |
| <i>Preliminary Analysis</i> | 35 |
| <i>General Analysis</i> | 36 |
| <i>Critical events mentioned during free recall and follow up.</i> | 36 |
| <i>Responding to focused questions about critical events.</i> | 36 |
| Discussion | 37 |
| <i>Limitations</i> | 43 |
| References | 45 |
| Appendix A | 50 |
| Appendix B | 53 |
| Appendix C | 54 |

Two children witness the same event. One of the witnesses is confident that he remembers all the details while the other witness is much less confident about what he remembers about the event. Which of the two witnesses would you find more credible? Research has shown that we prefer to believe a confident witness more than a less confident one (Penrod & Cutler, 1995). A number of studies have also shown that jurors in particular, find confident witnesses more credible than less confident witnesses (Cutler, Penrod, & Steve, 1988; Lindsay, Wells, & Rumpel, 1981; Wells, Lindsay, & Ferguson, 1979). The degree to which confidence match accuracy in our memory performance is called realism, where a perfect realism means being just as accurate as you are confident (Lichtenstein, Fischhoff, & Phillips, 1982). But, people who are more confident do not have to be more accurate. The relationship between accuracy and confidence is called calibration. A well calibrated witness is a person who is equally confident as he or she is accurate about information that they remember. This is also called meta-cognitive realism. In an experimental study, Allwood, Innes-Ker, Holmberg and Fredin (2008) found that 9-10 year-old children showed perfect realism when they could decide for themselves which information to volunteer as answers to open format questions. This means that children were just as confident as they were accurate when they were asked a general free recall question about an experienced event. But, do children continue to show perfect realism during follow up questions and focused questions?

Another interesting aspect is whether, when children witness the same event, there could be a potential influence on each other's testimonies. Now, we are going back to the example above. Two children witness the same event and then engage in a discussion about this event. Will this discussion affect their accuracy and confidence about what they believed happened?

These are very important questions regarding children's abilities in a witness situation that need to be investigated further within a scientific paradigm. Such research will not only help authorities and courts with the aim to judge the credibility of a child's testimony, but is also important to safe-guard the legal rights of children.

One aim of this study is to investigate in which way children's accuracy and confidence is effected by question format and peer discussion regarding a co-witness situation. Another aim is to examine the memory effect of the so called post-event information that a child may have obtained from another co-witness after engaging in a peer discussion regarding an event.

In an attempt to make the following exposition more comprehensible we will first present a review of the research regarding children's memory and the effect of question format. This will be followed by a summary of the effect of different types of questions on event recall and confidence judgments. We will continue with a summary of the research regarding suggestibility after which follows a review of the effect of peer discussions on the realism in confidence and a review of the effects of post-event information.

Children's Event Recall Memory

In a recent review by Goodman and Melinder (2007), the authors conclude that the memory consists of many complex interacting systems that affect humans' ability to encode, store and recall different types of information. The developments of these systems are highly dependent on the age and the individual maturity of the child. As children grow older their event recall becomes more informative, elaborative and complex (Pipe, Lamb, Orbach, & Esplin, 2004). There are also other factors that can affect event recall. One is the amount of time that has passed between the memory phase of encoding and the phase of retrieval . As a function of time memory recall becomes less accurate and less complete. Studies also suggest

that prior knowledge has a positive effect on memory recall (Greenhoot, 2000; Sutherland, Pipe, Schick, Murray, & Gobbo, 2003). Children that have more knowledge about a certain type of events remember more details about the event during recall than children with a lesser amount of knowledge. But, there still exists controversy in the scientific literature regarding children's memories of traumatic events (Pipe et al., 2004). While some researchers claim that stress improves children's accuracy about the remembered event, other researchers claim that stress contributes to a lower level of accuracy. It is thus unclear if there are different memory mechanisms that underlie the phases of encoding, storing and retrieval of traumatic memories than the mechanisms that underlie the processing of more neutral events and in what way stress affects memory. Although these are valid questions, it needs to be pointed out that far from all witnessed events that later becomes forensically important, have a traumatic impact on the child. For example, a man collects a child from the playground and another child witness the whole event. If the child has witnessed a nonviolent abduction, like this one, it is highly likely that he has not understood the significance of what he has witnessed. The child that witnessed the event probably believes that it is a parent or a relative has come to collect the other child from the playground. Therefore is it equally important to investigate non traumatic events as it is traumatic events.

Different Question Formats

Three common types of questions are used when interrogators interview witnesses for an event: open free recall questions, open follow up questions and focused questions (for examples of different interview procedures see Goodman & Melinder, 2007; Lindberg, Chapman, Samscock, Thomas & Lindberg, 2003). An example of an *open free recall question* is "Tell me everything you remember about the event!" These types of questions give the

witnesses the freedom to decide how much information that is to be included in the answer. When being asked this type of questions, the witness might exclude information that he or she is unsure of, or information that does not seem relevant to the witness. The second type of questions we refer to as: *follow up questions*. This label includes what is often referred to in the scientific literature as open-ended prompts. An example of an open-ended prompt is: “What happened then?” Another example is: “What were the different people wearing?” The latter example is more focused but still defined as an open-ended question. Follow up questions regularly follows an open free recall question and may prompt the witness to give more information about a certain part of an event. Even though follow up questions results in the witnesses having less freedom in how they choose to answer them than open questions, it has the advantage of prompting the witnesses to reveal relevant information that might otherwise not have been reported (Allwood et al., 2008). The third type of question is called focused questions. These questions could either have yes/ no- formats or be questions that demand a more specific and closed answer, for example: “What was the color of the car?”

A vast amount of studies have investigated how open free recall questions affect children’s event recall performance. Research has shown that children have a very low error rate when it comes to answering free recall questions (Gee & Pipe, 1995; Rudy & Goodman, 1991). But, compared to research done on adults, these studies have still rendered some mixed results. In a review by Pipe et al. (2004) the conclusion was that children are just as accurate at recalling an event as adults, but that they report significantly less information. In other words their reports are less complete in comparison to adults. Contradictory to these findings Allwood et al. (2008) found that 12-13 year-olds were significantly less accurate than both adults and 8-9 year-old children in their event recall but found a similar significant difference in their completeness between all age-groups, with adults showing the highest degree of completeness.

Research shows that children's accuracy level is markedly lower for focused questions than for free recall questions (Dickinson, Poole & Laimon, 2005). One reason for this could be that specific questions are more likely to prompt children to give information that they have already forgotten or never encoded to begin with. Another reason for the lower level of accuracy in children could be that children feel socially pressured to answer questions that they do not fully understand.

Research investigating the effect of follow up questions on children's level of accuracy in their recollections of events is very limited. We have only found one study by Poole and Lindsay (2001), where 3-8 year olds participated in a science demonstration and were later interviewed using five open-ended prompts. The results from this study showed that children volunteered new information when being prompted and that there was no decline in accuracy caused by the five open-ended prompts. The importance of follow up questions when interviewing children is still suggested in several studies due to the large amount of new information (almost half of children's narratives) that was provided during this phase of questioning (Lamb et al., 2003; Sternberg, Lamb, Orbach, Esplin, & Mitchell, 2001).

Metamemory, Confidence and Realism in Confidence

Pansky, Koriat and Goldsmith (2005) suggested another factor that may affect memory performance namely, *metamemory*. In this review metamemory was defined as: "what people know about their memories and how that knowledge is put to use in regulating what they report" (p. 94). Metamemory is a crucial part of a witness' competence. To illustrate this phenomenon, consider an eye-witness who is interviewed by a police officer. To be of help to the police officer, the witness needs to discriminate between correct memories and incorrect

memories that he or she might remember and he or she needs to report all the memories that are correct while omit those that are not.

As mentioned earlier, a good confidence in your testimony does not necessarily mean that you are accurate and vice versa. Earlier researchers in forensic psychology used the point-biserial correlation between accuracy and confidence as a measure of realism and in this literature it is only in recent years that calibration has proved to be a more accurate measure (Olsson & Juslin, 2002). In the present study two calibration measures were used, namely bias and slope (Yates, 1994). Bias (also called over-/underconfidence) is the degree of over- or underconfidence that an individual is showing in his or her memory performance. In the field of calibration research the most frequently seen bias is the overconfidence effect (McClelland & Bolger, 1994). Slope is a discrimination measure that tells you how well a person discriminates between confidence for correct and incorrect memories. A person that is very confident when he is accurate but much less confident when he is incorrect, shows good discrimination.

Not many studies within the field of eye-witness research have used calibration measures as an indicator of realism in adults and children's confidence for event recall. When using the question format of free recall Allwood, Ask and Granhag (2005a) found that the adults were well calibrated. Only one study has investigated the realism in children's confidence during free recall. This study (Allwood et al., 2008) found that children in the ages of 8-9 showed perfect realism in a free recall task while 12-13 year old children and adults showed some overconfidence for the same task. Interestingly enough, when the same groups responded to focused questions the adults showed almost perfect realism while both groups of children showed overconfidence. For the free recall task there was no difference in discrimination between the three groups but there was a lesser degree of discrimination for focused questions in the youngest children compared to the other two groups.

Earlier research, in contradiction to Allwood et al. (2008), shows that both adults and children show overconfidence when responding to focused questions (Allwood, Granhag, & Jonsson, 2006a; Allwood, Jonsson, & Granhag, 2005b; Allwood, Knutsson, & Granhag, 2006b; Juslin, Olsson, & Winman, 2000). The present authors know of no study that has investigated the realism of confidence in children's recall for follow up questions.

Confidence Judgments and Gender Differences

A moderate amount of research has investigated gender differences regarding the level of realism. This research has yielded mixed results. For adults, Jonsson and Allwood (2003) found that there was no significant difference between females' and males' realism of confidence judgment for a word test. Even though very little research has been done regarding differences for children's realism, in most of the studies that have been carried out, no significant gender differences have been found (Allwood et al., 2005b; Roebbers, Gelhaar, & Schneider, 2004; Roebbers & Howie, 2003). A notable exception to these findings is a study by Allwood et al., (2006a) that found significant gender differences. The results of this study showed that girls were significantly better calibrated than boys and that girls were significantly less confident than boys.

The Suggestibility of Children

As mentioned earlier, research has shown that after a crime has been committed, a majority of witnesses talk to each other about what they have experienced (Patterson & Kemp, 2006). A later study by Candel, Memon and Al-Harazi (2007) refers to a particular example, where an incorrect description of the perpetrator, in the case of the murder of the Swedish Foreign Minister Anna Lindh, stemmed from contagious contacts between witnesses.

This incorrect description spread between the witnesses and severely obstructed the search for the perpetrator.

The social influence between people in a witness situation is a part of a broader field called *suggestibility*. Children's suggestibility has been studied since the beginning of the 1900's. The studies have examined the suggestibility of children's testimony in an effort to determine whether they would be reliable as witnesses. Lately these studies have increased in numbers typically because of children's increasing presence in the courtrooms as single witnesses to sexual abuse (Ceci & Bruck, 1993; Pipe et al., 2004). According to its broadest definition, *suggestibility* concerns the degree to which children's encoding, storage, retrieval, and reporting of events are influenced by a range of social and psychological factors (Ceci & Bruck, 1993; Goodman & Melinder, 2007). This broad definition suggests that it is possible to incorporate information and still be aware of the difference between the information and the original event, and it does not have to imply that the memory of the event is impaired. This was shown in a study by Cohen and Harnick (1980) where 9-year olds, 12-year olds, and college students were compared with regards to their recollections of events from a film. They were faced with misinformation from an interrogator, but were later interviewed without the misinformation. The results showed that the younger children were less accurate than college students but the results also showed that their accuracy got significantly higher after a second interview that did not include any misinformation. One explanation for this result is that the children may know what actually took place but choose not to report it because of pressure from the interviewer/practitioner or parent/peer/perpetrator. This further shows the importance of this type of research which can help the interviewers give instructions and ask questions that optimizes the accuracy of witness reports and give the interviewers the knowledge they need to interpret the information and the credibility of the witness (Ceci & Bruck, 1993; Goodman & Melinder, 2007).

Previous research thus shows that children are suggestible and that they may be more suggestible the younger they are. However, this research field is not yet fully understood. According to two reviews of the field, there seem to be a consensus regarding the fact that suggestion is a reality when it comes to children as well as do adults (Ceci & Bruck, 1993; Goodman & Melinder, 2007; Pipe et al., 2004). This has meant that the focus of the research has shifted from simply examining whether children are suggestible to determining under what conditions, internal as well as external, that children are prone to suggestion.

The two reviews of the suggestibility of children show that children are as highly resistant to suggestion and as unlikely to lie as adult witnesses when it comes to acts perpetrated on their own bodies (Ceci & Bruck, 1993; Goodman & Melinder, 2007). But the reviews also show that children are described as having difficulty distinguishing reality from fantasy and that they are particularly susceptible to guidance by powerful authority figures, and therefore as potentially less reliable than adults. Another conclusion the authors draw is that children's memory reports are particularly affected by misleading details that are peripheral to the event and when people's actions are more ambiguous.

This study is a part of a research field of suggestibility referred to as *the effect of misinformation*. The effect of misinformation is, in the present study, defined as; the supplementary information that affects memory during the encoding phase that does not correspond to the original representation of the witnessed event (Loftus & Hoffman, 1989). Witnesses have been shown to be sensitive to the effects of misinformation in a vast amount of studies (Ceci & Bruck, 1993; Gabbert & Memon, 2006; Pipe, et al. 2004). The source of misinformation can be photographs of the suspects, an article in a newspaper or a conversation between two witnesses (Pipe et al, 2004). A study by Gabbert et al. (2006), show that additional, post event information can result in memory change and also in the creation of

new memories, by either the incorporation of new details not witnessed, and/or the omission of details previously reported.

Misinformation and Confidence in Children

Research on the effects of misinformation and particularly peer discussion on children has, as mentioned before, been scarce. Children have been presumed to be more suggestible than adults when it comes to the effects of misinformation due to peer discussion. But in fact, very few studies have actually been done (Ceci & Bruck, 1993; Goodman & Melinder, 2007).

In the present study we investigated how a discussion about a previously witnessed event affects confidence and realism in confidence in their answers to free recall, follow up and focused questions regarding the event. There have previously been suggestions of a link between realism in confidence and misinformation. A study by Gabbert et al. (2006) showed that the person who first mentioned the item in the peer discussion, was the most likely to influence the other person with respect to the credibility of this item. The study showed that the effect of misinformation due to peer discussion is a reality but also that there seem to be a link between confidence in memory and the effect of misinformation in peer discussion. The person who first answered the question in the peer discussion was likely to be the most confident about those details of the event.

A study by Allwood et al. (2005b) examined the effects of written feedback on 12-year-old children's confidence judgments. The source of the feedback was presented as either being from a peer or a teacher. The authors found no difference in confidence judgments in regards to what type of source that was presented. But when the children received confirmatory feedback they were significantly more overconfident than when they received disconfirmatory feedback. In a later study by Allwood et al. (2006b), the authors investigated

the effect of written confirmatory and disconfirmatory feedback on adults. The result showed that confirmatory feedback led to significantly higher levels of overconfidence than in the control group. No significant difference was found for disconfirmatory feedback. The limitations of these studies are that they did not examine the effect of peer discussion or even oral feedback, instead they investigated the effect of written feedback.

Peer Discussion and the Incorporation of Post-event Information

Peer discussion and the incorporation of post-event information have not been the subject of more than a couple of research papers. A study by Hjelmsäter, Granhag, Strömwall & Memon (2008) examined the extent to which social influence results in omission and commission errors in children. This is an important issue as omission errors could result in details being overlooked in police investigations. Children individually interacted with a stranger and were then later interviewed about this event. The children were given misinformation by the stranger who either denied the presence of an item that had actually been there or suggested an item that had not been present. The first main finding was that the children could be influenced to add a detail, but not to omit a detail from their reports. Another aim of this study was to examine how peripheral details versus central details functions as a mediating factor of social influence. The results showed a significant difference between central and peripheral details in that children in the control conditions made significantly more memory errors with respect to the peripheral details compared to the central details. Results indicate that judgments about the accuracy of children's testimony must consider the possibility of exposure to misinformation prior to formal interviews.

In one study by Meade, Roediger and Bergman (2001) the participants were asked to watch slides in pairs, depicting household scenes. After the stimuli had been presented, half of

the pairs were asked to complete a collaborative recall test. This task required the subject and the confederate to take turns recalling items from each scene. The results revealed that there was a significant effect of misinformation that stemmed from the partners reports during the joint tasks. Further, this effect was greater when the scenes were presented for less time (15 sec) and when the intruded item was more peripheral. In another study, Gabbert, Memon and Allan (2003) asked the participants (6-7 and 11-12 years old) to watch a video in pairs. The subjects were made to believe that they were watching the same video clip, but in fact there were two clips, each one lasting 1 min. 30 sec, shot from two different angles. Each film perspective contained unique details that were only seen by one of the pair members. Half of the pairs were then asked to discuss the event before each witness participated in a recall test. In the control condition dyads were not allowed to discuss the event prior to recall. The event was a short film of a girl entering an unoccupied university office to return a borrowed book. This study produced two main findings: All age groups displayed a memory conformity effect. More than 70% of the participants in the discussion condition erroneously recalled an item acquired during the discussion. These participants thus incorporated post event information provided by their conversation partner in their memory reports. This effect was significant in both free and cued recall. The second finding was that the effect of peer influence was stronger in the older children than in the younger ones for free recall but not for the focused questions. The older children also recalled fewer items in total.

This study was replicated in 2007 by Candel, Memon, and Al-Harazi. The result was comparable. More than 60% of the children recalled at least one detail from the video that only could have been seen by the co-witness, whereas 23% of the children recalled an unseen detail in the individual witness condition. Inferences from the witnessed details might be a reason why children in the individual witness condition reported details from the alternative video. Research has shown that children are particularly vulnerable to self-generated

misinformation (Holliday, 2003) In the study by Candle et al. (2007) there was among similar critical details, a glass filled with Coca-Cola in one of the clips. Children, who watched this clip, might have inferred a bottle of Coca-Cola from this seen detail. Another finding was that older children showed a stronger memory conformity effect in free recall than younger children did, but that this effect was absent for the focused questions. Accordingly there seems to be a misinformation effect due to the peer discussion. But there are several problems with these studies. The critical objects were all peripheral (as opposed to central) events. Examples are: the text on a book, the presence of a bottle and a girl wearing a cap. Another problem with these studies is that the participants in the co-witness condition were asked to complete this task with another witness by discussing the event together and then providing a single collaborative answer to every question asked. The subjects were thus made to agree upon a joint answer to every question. This is a highly unrealistic goal that seldom occurs in natural discussions about an event and it provides very little room for different recollections of the events in the film. Moreover, the participants did not really witness the same event in the study by Candel et al. (2007), because the film was not actually shot at the same time. Two films were made, one film after the other in an effort to film the same sequence twice, having the main actor repeat her performance and omitting certain details in every version. This cannot strictly be compared to a co-witness event.

The present authors have tried to improve on the previous studies on peer discussion by basing it on a more natural method where the films are shot at the same time, just as in the study by Gabbert et al. (2003), using a video that shows a number of different actions and interactions between the people in the film. In the present study peripheral objects, central objects and actions were shown and asked about. The participants were specifically told that they did not have to come to an agreement in the peer discussion and they were also

instructed that it was very important that they only reported what they themselves had witnessed in the film.

Even if most researchers agree that the effects of misinformation can destroy the credibility of a witness it is still unclear how co-witnesses influence each other before formal interviewing. Peer discussion could possibly help the witnesses in some cases. It may be that unrealistic methods might have blown the potential influence of peer discussion out of proportion. The present authors' prediction is that children are quite competent in distinguishing between what they actually have witnessed and what they have heard from a peer.

One of the aims of the present study is to examine the effect of misinformation between witnesses due to peer discussion and especially two subtypes of effects that are of interest. First, the effect of misinformation on children's confidence and realism in confidence for event recall is examined. Secondly the transfer of details unseen by one of the witnesses while being exposed to post-event information is studied. Of particular interest is the frequency of transferred unseen details and how confident children are about these reported unseen details.

Hypotheses

The first hypothesis concerns the free recall performance. Even though studying a slightly older age group, we believed that we, in accordance with Allwood et al. (2008), would find perfect realism in confidence for 9-11 year olds that were not subjected to a peer discussion about a witnessed event.

The second hypothesis concerned both the children's performance on free recall and follow up questions. We hypothesized that children would show the same level of confidence and a perfect realism of confidence for follow up questions and for free recall. The argument

for this hypothesis is that children still have a certain amount of freedom regarding what and how much they report.

The third hypothesis concerned gender differences in respect to confidence and calibration measures. Based on the previous discussion, we hypothesized that there will be no gender differences in regards to confidence levels and level of realism in confidence.

The fourth hypothesis concerned the effect of peer discussion on realism, with regards to their performance on free recall, follow up and focused questions. Based on the discussion above, we believed that children will overall show more overconfidence in their free recall, follow-up, and focused questions, when they are subjected to a peer discussion about the event, compared to a neutral discussion.

The fifth hypothesis predicted that there would be a transfer of memories between children. However, we predicted that this difference would be smaller than has been shown in previous studies. This hypothesis is based on the discrepancies in earlier findings and research methods that includes joint tasks that forces the children to come to an agreement regarding the event and the lack of instructions before the interview that include only to report what they have witnessed.

Method

Participants

Originally 78 children were tested but for seven children data could not fully be obtained or could not be used because of various reasons, such as missing the second testing, not following the instructions given etc. Data was collected from 71 children, 11 9-yearolds, (7 girls and 4 boys), 51 10-yearolds (25 girls and 26 boys) and 9 11-yearolds (3 girls and 6 boys). In total 35 girls and 36 boys participated. For free recall and follow up questions valid

data for both question formats was collected from 68 participants (experiment $n = 33$, control: $n = 35$). Data from two of the participants from the original sample could not be collected due to technical error and one participant had no follow up-data. For focused questions valid data was collected from 70 participants (experiment: $n = 34$, control: $n = 36$). One participant's data was excluded due to having more than five missing values and therefore were considered unreliable. The participants were recruited from third and fourth graders in five different schools in the south of Sweden. The participation rates of children in the classes that were offered to participate ranged from 85-90%. Three schools were situated in rural areas and two schools were located in smaller towns. The children were mostly of middle-class socioeconomic background. The parents and the children were asked to sign a paper that stated that their participation was voluntary (see Appendix A).

Design

Since different confidence rating scale was used for the three question formats, free recall and follow up questions were separately analyzed from focused questions. Data about misinformation was also analyzed separately. This division yielded three experimental designs. First, a factorial design 2×2 with the between-subject factor of discussion type (film discussion vs. neutral discussion) and the within-subject factor of question format (free recall question vs. follow up questions). Second, the focused questions were analyzed in an independent group design. Third, the data about misinformation was analyzed in a dependent and an independent group design.

Material

Video Clip

The stimulus consisted of a 8,5-min amateur film of a picnic. The video had no sound. The film was shot simultaneously, from two different angles, rendering two films, Film A and Film B. It shows three people, two women and one man, having a picnic on the grass in front of some trees. During the picnic two people in turn block the cameras partially. This means that when one camera is partly blocked the other camera is the only one that has a full view of the picnic. Most of the time, both cameras had a full view of the picnic. The details that are blocked are specific events: a hug, a high five, somebody takes a picture, somebody gets a phone call, somebody drops a piece of paper, and somebody receives a present. Three specific events are blocked in Film A, while three others are blocked in Film B. Thus for each child, regardless of which film version (A or B) he or she saw, there were three critical events that her or she couldn't have seen because the cameras were blocked. The critical events were evenly spread between the two films. The two films simulate the points of view of two potential witnesses situated at slightly different viewpoints observing the same event.

Equipment

The two film versions were presented on two similar laptops with high-resolution computer screens, in order for two children to be able to watch a version each of the film (A or B) at the same time. The interviews were recorded with an mp3-recording device or a mini-disc-recording device.

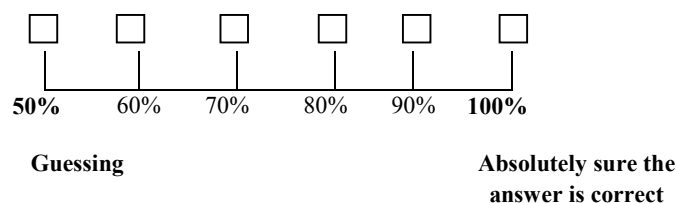
Questions used in the Discussion Conditions

The children, who were assigned to the experiment group, were asked to discuss what happened in the film using a set of open and focused questions. The open question was stated as follows (translated from Swedish): *Discuss with each other what the different people in the film did.* The focused question was stated as follows: *Discuss the following sentences and discuss which of the events took place in the film.* An example of a sentence was: *Someone took a picture?* Four of the sentences concerned events that were visible in both film versions. Two of the sentences concerned events that only film A had shown and two sentences concerned two events that only film B had shown. Finally two sentences concerned events that were not present in neither of the film versions (see Appendix B). The participants were specifically told that they didn't have to come to an agreement regarding the events in the film. In the control group, the children were asked to discuss a set of questions regarding neutral subjects (see Appendix C).

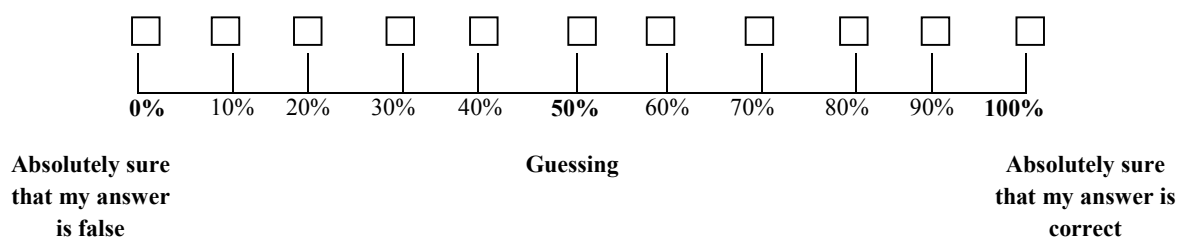
Questionnaire for the Focused Questions and Confidence Scales

A questionnaire with focused questions was used. The questionnaire consisted of thirty-one two-alternative questions about the content of the film. The questions concerned details about the actors' appearances and about some of the events that took place in the film. There were also questions about four of the critical events, two that only one or the other film showed, and twenty-nine details that were present in both films. The confidence judgment for each of the two-alternative questions was made on a 6-point scale ranging from 50% (guessing) to 100% (absolutely sure the answer is correct) (See Figure 1). Two examples of the focused questions (with answer alternatives) are; *Of what color was the picnic blanket? (a) blue, (b) red,* and *In the middle of the film, a guy arrives with a ball. He receives something from one of the girls. What? (a) a present (b) a flower.* For free recall and follow

up questions a confidence scale (0-100) was placed under each statement. The confidence judgment was made on an 11-point scale, ranging from 0% (absolutely sure that my answer is false) to 100% (absolutely sure my answer is correct). 50% meant that they were guessing (See Figure 2).



Figur 1. Confidence scale used for focused questions.



Figur 2. Confidence scale used for free recall and follow up questions.

Preparation of the Free Recall and Follow Up Material for Confidence Judgments

For every participant an individual memory-questionnaire was developed based on their own testimonies in the interview section. In order to allow for the participants to rate their confidence in each part of their free recall responses and for their responses to the follow up questions, the interview transcript was divided into low-level units corresponding to the participants' recall during their interviews. The coding principles employed by Allwood, Ask et al. (2005) were used to segment the free recall protocols into unit-statements.

Procedure

Every participant was, within his or her school, randomly assigned to an experiment condition and a control condition, with the constraint that there be an equal number of children from each gender in each condition. The experiment consisted of two separate sessions.

Session One

Session one took place at the children's school. The experimenter first briefly explained how the experiment would be conducted. Then, the children were called to a separate room in pairs. The order by which they were called in was randomized within the five schools. The children were told they were going to watch a film about a picnic. They were also told to pay close attention to the film. The pairs were placed in a small room and the films were screened on two different computers, simultaneously. The participants were made to believe that they were shown the same film. The computers were placed in such a way that it would be impossible to watch each other's films.

Then half of the pairs which were assigned to the experiment condition were asked to discuss what happened in the film using a set of open and focused questions. The discussion was supervised to make sure that all of the questions were discussed. The other half of the pairs were assigned to the control condition. They were then asked to discuss their favorite subjects in school and their favorite pastime activities. These discussions were also held using a set of open and focused questions.

After the discussions, pairs in both conditions performed a short filler-task for five minutes which consisted of a children's crossword puzzle. Then, the pairs were separated and interviewed by the experimenters using a free recall question and three follow-up questions. A

manuscript was used to standardize the interview. The interview part of the study took place simultaneously for the two pair members, in two separate rooms. Which one of the two experimenters that interviewed a specific participant was randomized within the pairs. The interviewer first established a rapport with the child by asking how he or she was doing. When the interview began the participants were asked to report everything that they remembered in the film. In this context they were asked only to report things they themselves remembered from the film. They were then asked three open follow-up questions as follows. First, they were asked to report if they remembered any more events taking place in the film. Then they were asked what they remembered about the setting/environment and finally they were asked what the people in the film looked like. After the interview the participants were instructed not to discuss the experiment with anybody for a period of a week until they had completed the second part of the experiment.

Session Two

Session two took place approximately one week after the first session. The participants first received a fifteen minute long training session concerning probability assessments. First they were instructed on how to fill in the joint free recall and follow up questionnaire. An example of how the statements would be like was given, for example: *There was a brown dog.* The testing instructor then demonstrated different confidence ratings according to how confident she was regarding the question being asked. For example she could say: *If I'm absolutely sure that the dog was black I would mark the 100 % box. But maybe I've changed my mind from last week and I'm now certain that the dog was actually brown . Then I would mark the 0% box, because I'm absolutely certain that my answer is false.* After a couple of examples like the ones given above, the children were given oral examples of different degrees of confidence that they themselves should decide how to

confidence rate. The instructor could for example ask: *“If I’m guessing that the dog was black, it might just as well have been brown, which box would I mark then?”*

For the focused question questionnaire a similar demonstration followed. But this time the children were instructed to first pick an answer out of two alternatives and then confidence rate how sure they were that they had answered correctly. The children were told to answer all questions and that it was ok to guess since they could tell the experimenters they were guessing when they marked how confident they were. The training session was standardized to make sure that every participant got the same information. We also tried to answer the participants’ questions during this session. This was done to make sure that the participants fully understood the probability scale used when making the confidence judgments.

Then, they were handed two questionnaires. The questionnaire answered first contained the participants’ own answers from the interviews (i.e., including statements from both their free recall and answers to the follow up questions). For each of statements, they were asked to review their own statements and to make a confidence judgment. Next, they were given a questionnaire containing two-alternative focused questions and the participants were instructed to choose one of the alternatives. Directly after each of their answers they were instructed to make a confidence judgment regarding how sure they were that they had answered correctly. During the confidence judgments the experimenters were available to the participants to answer any questions about confidence judgments and to make sure the questionnaire was filled out properly. When the participants had completed both of the questionnaires they were thanked and excused.

Scoring Procedure and Coding

All statements in a child's response to the free recall and to the follow up questions were scored as either correct or incorrect based on the content of the film. Statements that could not be objectively verified, for example psychological states, were not scored. Scoring rules were set up for ambiguous cases. For example, if the child described one of the girl's blond hair as being "white" or "bright yellow" it would be considered a correct statement.

The focused questions were simply scored by checking if the person had chosen the correct alternative to the question. If a child forgot to answer a question or marked both alternatives to one question as correct these questions were considered missing values. If a questionnaire had more than five missing values it was regarded as unreliable and excluded from the analysis.

Calibration Measures

To measure the degree of realism in confidence two calibration measures were calculated. The first calibration measure, *bias* indicates if a person shows over- or underconfidence in his or her judgment. Bias is calculated by taking a person's average confidence judgments minus the same person's average proportion of correct answers. A positive value indicates that the person is overconfident, in other words, more confident than he or she is accurate. A negative value on the other hand indicates that a person is underconfident, he or she is more accurate than he or she is confident. A value near zero, indicates that a person shows good realism in confidence, in other words, the person is well calibrated (Yates, 1994).

The other measure is a discrimination measure called *slope*. Slope indicates how well a person discriminates in confidence between correct and incorrect answers. Slope is calculated by taking a person's overall mean confidence rating for incorrect items and subtracting it

from the same person's overall mean judgment for correct items. Poor discrimination yields a number near zero while good discrimination (if the scale ranges from 0-100) yields a number near one (Yates, 1994).

Results

Analyzes of Peer Discussion and Questions Formats

Preliminary Analysis

Preliminary analysis did not uncover differences that could be related to the different film perspectives nor did it reveal any effect of interviewer. No gender effects were found for the free recall and follow up questions. However, gender effects were found for three of the measures in focused questions. A significant gender effect was found for accuracy for the focused questions, $F(1,68) = 6.531$, $p = .013$, partial $\eta^2 = .088$, such that the girls ($M= 69.7\%$, $SD= 9.2\%$) were more accurate than the boys ($M= 64.2\%$, $SD= 9.3\%$) when answering the focused questions. Girls were also significantly less confident ($M= 75.6\%$, $SD= 10.1\%$) than the boys ($M= 81.2\%$, $SD= 9.8\%$) when responding to the focused questions, $F(1,68) = 5.488$, $p = .022$, partial $\eta^2 = .075$. Boys were also significantly more overconfident ($M= 0.169$, $SD= .121$) than the girls ($M= 0.059$, $SD= .119$), $F(1,68) = 14.790$, $p < .0005$, partial $\eta^2 = .179$.

In Table 1 the means and standard deviations for accuracy, confidence, bias and slope for each condition (event discussion vs. neutral discussion) and each question format (free recall, follow up questions, focused question) is displayed. Completeness, the number of statements recalled, was also calculated for the two question formats free recall and follow up.

Table 1. Mean values, SDs, and n for the two conditions and three questions formats for the different dependent measures

| | Film Discussion | | | Neutral Discussion | | |
|---------------------------|-----------------|---------|-----|--------------------|---------|-----|
| | M | SD | n | M | SD | n |
| <i>Free recall</i> | | | | | | |
| No. of statements | 13.8 | (6.6) | 33 | 14.8 | (8.5) | 36 |
| Accuracy | 96.0% | (5.9%) | 33 | 96.6% | (7.0%) | 36 |
| Confidence | 96.7% | (4.5%) | 33 | 96.7% | (5.7%) | 36 |
| Bias | 0.008 | (0.062) | 33 | 0.003 | (0.083) | 36 |
| Slope | 0.245 | (0.370) | 13 | 0.114 | (0.208) | 9 |
| <i>Follow up question</i> | | | | | | |
| No. of statements | 10.5 | (4.6) | 33 | 11.0 | (3.7) | 35 |
| Accuracy | 89.7% | (10.6%) | 33 | 89.0% | (10.5%) | 35 |
| Confidence | 92.7% | (8.6%) | 33 | 93.4% | (8.7%) | 35 |
| Bias | 0.030 | (0.127) | 33 | 0.044 | (0.121) | 35 |
| Slope | 0.186 | (0.293) | 21 | 0.120 | (0.184) | 22 |
| <i>Focused questions</i> | | | | | | |
| Accuracy | 67.8% | (8.6%) | 34 | 66.2% | (9.7%) | 36 |
| Confidence | 78.7% | (10.3%) | 34 | 78.2% | (10.4%) | 36 |
| Bias | 0.107 | (0.121) | 34 | 0.119 | (0.142) | 36 |
| Slope | 0.121 | (0.067) | 34 | 0.111 | (0.081) | 36 |

General Analysis

A mixed ANOVA 2×2 was computed for the two conditions (film discussion vs. neutral discussion) and for two of the question formats (free recall, follow up questions) for each of the dependent measures (completeness, accuracy, confidence, bias, and slope). The focused questions were analyzed separately and in an attempt to partial out the gender effect an ANCOVA was carried out on the dependent variables accuracy, confidence, bias and slope.

Completeness. Analysis of number of statements recalled showed that there was a main effect of question format, $F(1,66) = 14.37, p < .0005$, partial $\eta^2 = .179$. The children recalled a higher number of statements during free recall than during follow up. No main effect was found for condition, $F(1,66) = 0.117, p = .734$, partial $\eta^2 = .002$.

Accuracy. There was no main effect found for accuracy between children who discussed the event with their peers and the children who discussed a neutral event, $F(1,66) = 0.000, p = .991$, partial $\eta^2 = .000$. However there was a main effect for question format, $F(1,66) = 24.266, p < .0005$, partial $\eta^2 = .269$. Children showed a higher degree of accuracy during free recall than during follow up. After adjusting for gender effects, by means of an ANCOVA, no significant difference between the two conditions was found for accuracy during focused questions, $F(1,67) = .530, p = .469$, partial $\eta^2 = .008$.

Confidence. There was a main effect for questions format, $F(1,66) = 10.762, p = .002$, partial $\eta^2 = .140$, showing that children's confidence significantly declined when they responded to follow up questions. No main effect was found between the two conditions, $F(1,66) = 10.762, p < .002$, partial $\eta^2 = .140$. When adjusting for gender effects, no

significant difference was found between the two conditions for focused questions, $F(1,67) = .038, p = .845, \text{partial } \eta^2 = .001$.

Calibration curves. For free recall, calibration curves are provided in Figure 3, and for follow up questions in Figure 4 and for focused questions in Figure 5. Each figure contains two calibration curves, one for the film discussion condition and one for the neutral discussion condition. On each calibration curve, the number of times each confidence class was used is marked with a tag indicating the proportion percentage. It should be noted that for free recall and follow up questions a majority of the confidence ratings made (between 73-92%) were a 100% ratings. During focused questions, however, the answers were more evenly distributed on the 50%-100% scale. A curve that is above the reference line indicates underconfidence and a curve above the line indicates overconfidence. Even though the

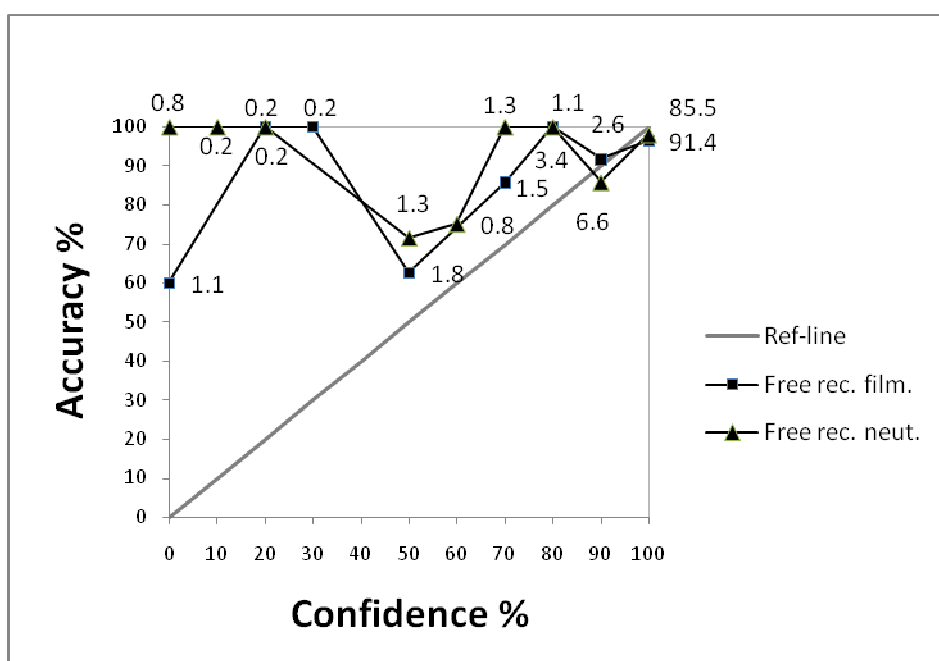


Figure 3. Calibration curves for free recall for both the film discussion condition and the neutral discussion condition.

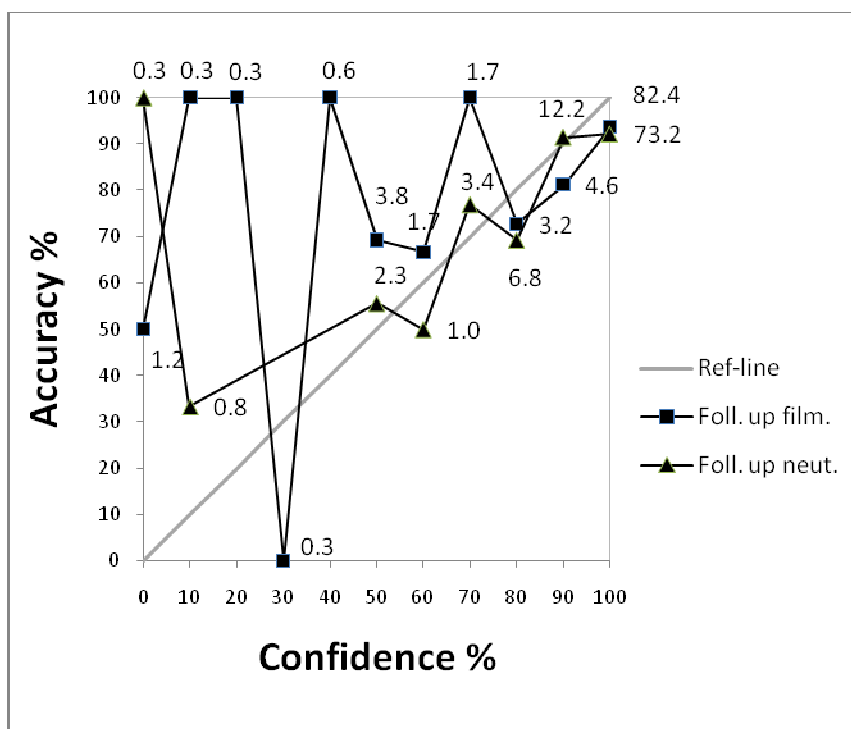


Figure 4. Calibration curves for follow up questions for both the film discussion condition and the neutral discussion condition.

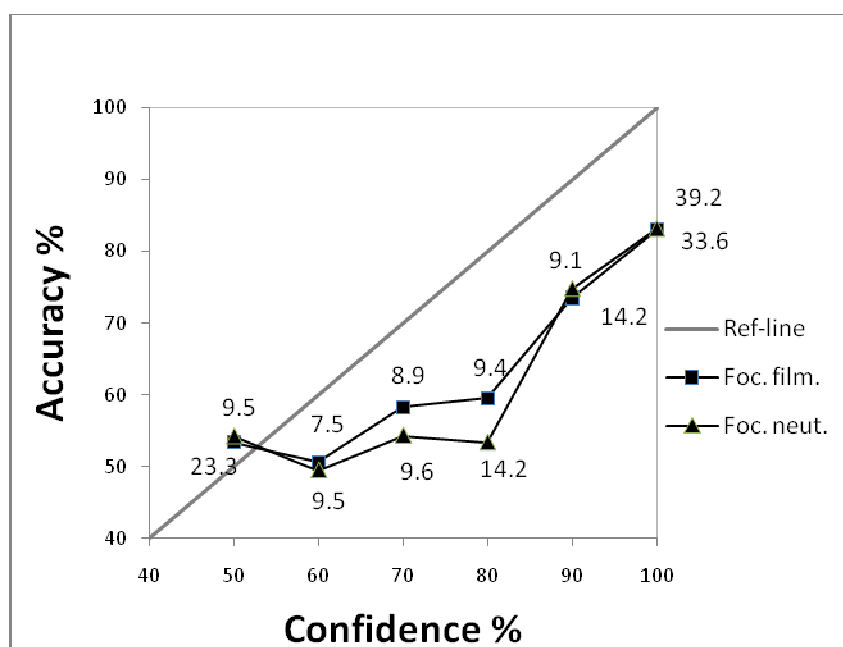


Figure 5. Calibration curves for focused questions for both the film discussion condition and the neutral discussion condition.

majority of the visual curve in free recall and follow up (Figure 3 and Figure 4) is above the reference line the majority of the answers, as already mentioned lies on the 100 % confidence level and slightly below the reference line. When looking at the focused question calibration curve a very visible overconfidence trend can be spotted. The curve is mostly below the reference line. This indicates that that the children were less accurate than confident during free recall for all confidence classes, except for the 50% confidence class.

Bias. The children were significantly less overconfident during free recall, $F(1,66) = 4.408, p = .040$, partial $\eta^2 = .063$, thus the overconfidence significantly increased during the follow up questions. There was, however, no significant difference in overconfidence between children who discussed the event and children who discussed a neutral event during free recall and follow up questions, $F(1,66) = .423, p = .517$, partial $\eta^2 = .006$. No significant effect was found between the different conditions for focused questions, $F(1,67) = .038, p = .712$, partial $\eta^2 = .002$, when controlling for gender effect. Since a bias level not significantly different from 0 indicates perfect realism, a t-test was calculated for the different questions formats; free recall and follow up questions. The null hypothesis was that there would be no difference between the measured bias and the value 0. For free recall, no significant difference was found, $t = .561, df = 68, p = .577$, two-tailed, thus children showed perfect realism when responding to the free recall question. For the follow up questions a significant difference was found, $t = 2.494, df = 67, p = .015$ two-tailed. Since there might be a gender difference, two t-tests was computed separately for boys' and girls' bias measures for follow up questions. The boys bias measure ($M = 0.059$) differed significantly from 0, $t = 3.647, df = 33, p = .001$ two-tailed, and therefore showed a slight overconfidence. However, the girls continued to show perfect realism during follow up questions, $t = .605, df = 33, p = .549$ two-tailed, since their bias measure ($M = 0.015$) did not differ significantly from 0.

Slope. Since slope is a discrimination measure that indicates how well a person discriminates between correct and incorrect answers by means of their confidence, it can only be calculated for individuals who make error in their answers. The analysis of the discrimination measure rendered no main effect for question format (free recall and follow up), $F(1,13) = .239, p = .633, \text{partial } \eta^2 = .018$, nor for the different conditions, $F(1,13) = .023, p = .882, \text{partial } \eta^2 = .002$. For focused questions there was no effect of condition on slope, $F(1,67) = .338, p = .563, \text{partial } \eta^2 = .005$.

Analyzes of Critical Events and Question Format

Preliminary Analysis

As noted, the interviewers checked that all the couples that were assigned to the condition of discussing the film actually did discuss critical events during the experiment, that is events that only one of the children could have seen. During the interview, when responding to free recall and follow up questions, 28% ($N = 11$) of the children that were asked to discuss the film said that they recalled a critical event, that they could not have seen. One child in the neutral discussion condition also claimed to remember a critical event during the interview. Because of the small sample size a non parametric test for repeated measures, Friedman's, was used.

For focused questions regarding critical events (all in all two for each subject) two measures was calculated: proportion of correct answers and confidence. The data did not meet the normality assumption. Therefore, when checking for gender effects the data was collapsed between the two conditions and the non-parametric Mann-Whitney test was used. The analysis yielded a significant difference in confidence ratings between genders for the critical events during focused questions, $U = 408.000, N_1 = 35, N_2 = 34, p = .023$, two-tailed. Boys

were more confident ($M = 78.4\%$, $SD = 17.6\%$) that they had seen the critical event than girls ($M = 69.1\%$, $SD = 17.0\%$). Since the number of girls and boys were equally distributed in both conditions (film discussion: girls $n = 16$, boys $n = 17$, neutral discussion: girls $n = 18$, boys $n = 18$) an analysis was carried out for the entire data without controlling for gender, after which the data was separated and an analysis of gender differences for the two conditions was performed.

General Analysis

Critical events mentioned during free recall and follow up. To investigate if there were any differences in confidence within subjects for the critical events and other events mentioned during free recall and follow up, a Friedman's test was computed. There was no significant difference between the confidence ratings for critical events and other information reported during free recall and follow up questions, $\chi^2(2, N=12) = .229, p = .892$.

Responding to focused questions about critical events. A Mann-Whitney was computed for the proportion of correct answers for the between-subjects conditions for the two focused questions relating to the critical events. Analysis revealed that children that were allowed to discuss the film ($M = 69.8\%$, $SD = 37.8\%$) choose the correct answer more often than the children who discussed a neutral subject ($M = 50.0\%$, $SD = 37.4\%$), $U = 425.000, N_1 = 33, N_2 = 36, p = .029$, two-tailed.

A Mann-Whitney test was also computed to investigate if there were any differences in confidence ratings for the critical events with regards to condition. There was a significant difference in confidence rating for children who discussed the film ($M = 79.2\%$, $SD = 16.0\%$) and children who discussed a neutral subject ($M = 68.9\%$, $SD = 18.2\%$), $U = 403.500, N_1 =$

33, $N_2 = 36$, $p = .021$, two-tailed. Children who discussed the film were more confident about the critical events than children who had discussed a neutral subject.

A Wilcoxon signed ranks test was also computed for the children who discussed the film to investigate the differences in confidence ratings for critical events compared with their confidence ratings on the other focused questions. Analysis revealed that there was no difference in confidence for the critical events questions ($M = 79.2\%$, $SD = 16.0\%$) and the other questions ($M = 78.7\%$, $SD = 10.3\%$), $z = -1.197$, $N - \text{Ties} = 33$, $p = .231$. The children who had a neutral discussion were significantly less confident for the critical events ($M = 68.9\%$, $SD = 18.2\%$), than for the other focused questions ($M = 78.2\%$, $SD = 10.4\%$), $z = -3.504$, $N - \text{Ties} = 36$, $p < .0005$.

Four Mann-Whitney's were computed for gender differences in the different conditions. Girls who discussed the film chose the correct answer as often ($M = 62.5\%$, $SD = 38.7\%$) as the boys ($M = 76.5\%$, $SD = 35.9\%$), $U = 107.500$, $N_1 = 17$, $N_2 = 16$, $p = .309$, two-tailed. But girls were significantly less confident ($M = 72.1\%$, $SD = 16.2\%$) about focused questions regarding critical events than the boys were ($M = 85.9\%$, $SD = 13.0\%$), $U = 70.500$, $N_1 = 17$, $N_2 = 16$, $p = .017$, two-tailed. No gender differences were found in accuracy or confidence ratings for focused questions regarding critical events for children who discussed a neutral subject.

Discussion

This study investigated 9-11 year old children who watched a film from one of two perspectives (A & B) of the same event. Half of the children were then instructed to discuss the film in pairs and the other half of the children discussed a neutral subject in pairs. The children were later interviewed using three question formats namely, free recall, follow up

questions and focused questions. The first aim of this study was to investigate the impact that peer discussion and question format have on the realism in confidence in children's testimonies. The second aim was to investigate if peer discussion leads to children reporting events that they cannot have witnessed and the confidence they show for these critical events. The results of this study will now be discussed in relation to the earlier stated hypotheses.

The first hypothesis regarding free recall was confirmed since the result shows that children in the age group of 9-11 showed perfect realism in confidence. They were just as accurate in their free recall testimonies as they were confident. This result replicated the result found in Allwood et al. (2008) and is quite remarkable since no other age-group in earlier studies has proved to show perfect realism. The reason that children in the age group of 9-11 shows perfect realism in confidence seems to be that they choose to only report information that they are highly confident is accurate. Our results support this explanation since a lot of the children made no errors at all during their free recall report. The calibration curve illustrated in Figure 3 also shows that the majority of the statements during free recall were statements that they one week later scored to be 100 % confident about. The result therefore seems to indicate that children in the age range of 9-11 years old almost only report information that they believe is accurate and that they are highly accurate during free recall report.

The second hypothesis, stated that there would be no difference in level of confidence and realism in confidence between the two question formats free recall and follow up, could not be confirmed. This study was not able to replicate the findings of Poole and Lindsay (2001), who found the same level of accuracy for follow up questions as for free recall in children's testimonies. Instead this study found that children were significantly less accurate during follow up questions than during free recall. One reason could be that the amount of

freedom to report what they remembered decreased with the narrower questions during follow up. This might have led to children feeling more forced to report things they were not a 100 percent sure of. This would also explain why children were significantly less confident during the follow up questions. However, the results indicate that children's confidence did not decrease as much as their accuracy during the follow up questions; the bias measure shows that the children went from being perfectly realistic during free recall to being slightly overconfident during follow up questions. The results from hypothesis one and two suggest that children are highly competent witnesses. Children's reports are highly accurate when they are being interviewed about a witnessed event with free recall and follow up questions. This proves the importance of interrogators using free recall questions and open follow up questions as much as possible. The results also indicate that children are competent witnesses when it comes to assessing their confidence regarding their memories during free recall questions and open follow up questions. These results suggest the importance of interrogators investigating not only the statements of a witness but also the witness' confidence about his or her statements not only in free recall but also when it comes to follow up questions. Another implication for the practitioner is that the child's confidence in his or her memory report, which has been acquired through free recall questions, can be used as a tool to judge the credibility of this part of the witness report.

Our third hypothesis stated that there would be no gender differences in confidence and realism in confidence was confirmed for free recall and follow up questions. However, significant gender differences were found during focused questions. Girls were more accurate but also significantly less confident than boys during focused questions. Boys were also slightly more overconfident during the same question format. These differences are quite interesting and suggest that girls are more competent than boys at answering focused questions. Are the girls in this age group cognitively more mature than the boys or could this

discrepancy have anything to do with differences in up-bringing? One explanation could be that the environment in which girls grow up teaches them to be more self-critical than boys. It could be that when growing up, girls are more often told to question themselves and their own judgment and therefore are more realistic in judging their mental abilities. These results suggest that girls are more competent when being interviewed with focused questions. And interrogators should keep in mind that even though girls are less confident than boys they report more accurate information. During the analyses of bias measure for follow up question we found when separating the group by gender that girls continued to show perfect realism during follow up questions. But, it should be noted that the difference between girls' and boys' bias measures were so small that it can hardly be considered a gender difference.

The fourth hypothesis stated there would be a difference in confidence and realism of confidence as an effect of peer discussion. This hypothesis could not be confirmed. Peer discussion seemed to have no effect on any of the measurements for the different question formats. Since a lot of the information rendered from the film was the same for the children even though the perspectives were slightly different, we believe that the discussion of the film would mostly have lead to confirmatory feedback rather than to disconfirmatory feedback. As mentioned earlier, studies have shown that confirmatory feedback led to the participants being significantly more confident and being significantly more overconfident while disconfirmatory feedback had little or no effect (Allwood et al., 2005b; Allwood et al., 2006b). Surprisingly enough, in the current study, although the pairs that discussed the film were submitted to confirmatory feedback it did not affect their confidence or their realism in confidence. One important difference between the current study and the studies by Allwood et al. (2005b) and Allwood et al. (2006b) is that the latter studies did not investigate the effects of peer discussion, but instead the effect of written confirmatory feedback. Maybe, confirmatory feedback as a result of peer discussion has less of an impact on confidence and

realism in confidence than other types of confirmatory feedback. These results indicate that a child witness is not more confident after talking with a peer co-witness than a child witness that has not been subjected to a peer discussion. So when assessing a child's confidence, the interrogator need not to worry that the child's confidence about his or her memories might have been comprised by peer discussion.

The fifth hypothesis predicted that there would be a transfer of memories between children as an effect of peer discussion. This hypothesis was confirmed. During the interview, when responding to recall and follow up questions, 28% (n = 11) of the children that were allowed to discuss the film said they recalled a critical event, that they could not have seen.. One of the children who did not discuss the witnessed event but instead discussed a neutral event, claimed to remember an event only witnessed by the peer during the interview. This memory report could thus not have stemmed from the discussion. The most likely reason for the child in the neutral discussion group to report a critical event is that he randomly reported an event that he thought had been in the film. The fact that this actually was a real event in the film that his classmate saw could have been a coincidence. On occasions, although very seldom, it happened that a child reported events that were not present in either film perspectives. The percentage of children, in the film discussion group, reporting critical events were thus a lot lower in the present study than in the studies of Gabbert et al. (2003) and Memon et.al. (2007). Memon et al. (2007) found that over 60% of the participants reported an unseen detail after discussing the target event, whereas in Gabbert et al.'s study over 70% of the participants reported such a detail. An explanation for our finding might be that the interviewers in the present study explicitly told the participant only to report what they themselves had seen in the video clip, which might prompt the participants to consciously evaluate their memories for the event. Another explanation might be that the peer discussion didn't involve a forced agreement regarding the events in the video clip which has

been the norm in previous studies. The results from hypothesis four and five indicates that even though children's confidence is not effected by peer discussion there is the effect of some children reporting events that only their co-witness saw. This suggests that when possible, during a criminal investigation, children should be instructed not to talk about the event with a co-witness. This can be difficult to achieve since it is very likely that the discussion among co-witnesses already has taken place when the police arrives at a scene, rendering the peer discussion an already established fact. However, we suggest that interrogators can minimize the effects of peer discussion in the interview situation. Our results indicate, contrary to earlier studies, that a much smaller proportion of children who are subjected to a peer discussion about a witnessed event, report things that they cannot have seen themselves. When given clear instructions child witnesses are much more competent at sorting out what they themselves have seen from other sources of information. Therefore our results suggest that it is important for interrogators to make it clear to the child witnesses that they should only report what they themselves have seen. This could easily be achieved with an instruction at the beginning of the interview. By doing this, a much smaller percentage of children will report things that they have not seen themselves.

We found no significant differences in confidence ratings for the critical events mentioned during free recall, follow up questions and other statements made during the same phase. Since this analysis was done on such a small sample ($n = 11$) there can hardly be any conclusions drawn, although the results seem to indicate that the children didn't discriminate well between transferred memories and their own genuine memories

When responding to the two focused questions relating to the critical events, analysis revealed that, children in the film discussion condition chose the correct answer more often than the children who discussed a neutral subject. Children who discussed the film were also more confident about the critical events than children who discussed a neutral subject. It is not

so surprising that the children in the film discussion condition chose the correct answer more often since their co-witness had mentioned the critical events. However, the fact that the children who discussed the film were much more confident than the children who discussed a neutral subject proved that the children became more confident after the peer discussion where the critical events took place even though they never actually witnessed them. However, it is difficult to draw any conclusions from these findings. A child might well be more confident that an event happened if a friend told him about it and still know that this event was something that he did not see himself.

When we separated the children who discussed the witnessed event by gender we found a difference between the boys and the girls confidence ratings of critical events, when they were responding to focused questions. Even though there was no difference in frequency between the boys and the girls when choosing the correct answer, the girls were significantly less confident regarding the critical events compared to the boys. However, separating the film discussion condition by gender rendered a smaller sample for analyses which makes any conclusions, drawn for this analysis, highly uncertain.

Limitations

One limitation of our study is that the final analysis of confidence for critical events mentioned during free recall and follow up had to be conducted on a very small subsample of only eleven individuals. Therefore any conclusions that might be drawn from this data are doubtful and must be scrutinized.

Another limitation is that we do not know which type of feedback the peer discussion consisted of. If the discussion had been recorded, an analysis of the content could have been done to establish if confirmatory feedback actually was the more common type of feedback.

Since different types of written feedback seem to lead to different effects on confidence judgments (Allwood et al., 2006b) it's possible to assume that different types of oral feedback effects confidence judgments differently.

Finally, it is possible that the two co-witness prior relationship might have affected the influence of the peer discussion. Since the children in this study were randomized they could end up discussing with a classmate they did not particularly like or someone who they considered a dear friend. It is possible that these differences in relationship could have led to different effects during peer discussion.

Therefore more research with a larger subsample of individuals who report misinformation is needed. In the future, when investigating peer discussion, there is a need to control which type of feedback that is most common during the peer discussion and in what way different types of feedback affect peer discussion. The impact of the different relationships among children on the effect of peer discussion also needs to be examined further.

References

- Allwood, C.M., Ask, K., & Granhag, P.A. (2005a). The Cognitive Interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime & Law, 11*, 183-198.
- Allwood, C.M., Granhag, P.A., & Jonsson, A.C. (2006a). Child witnesses' metamemory realism. *Scandinavian Journal of Psychology, 47*, 461-470.
- Allwood, C.M., Innes-Ker, Å., Holmgren, J., & Fredin, G. (2008). Children's and adults' realism in their event-recall confidence in responses to free recall and focussed question. *Psychology, Crime & Law, 14*, 529-547.
- Allwood, C.M., Jonsson, A.C., & Granhag, P.A. (2005b). The effects and source of type of feedback on child witnesses' metamemory accuracy *Applied Cognitive Psychology, 19*, 331-334.
- Allwood, C.M., Knutsson, J., & Granhag, P.A. (2006b). Eyewitness under influence: How feedback affects realism in confidence. *Psychology, Crime & Law, 12*, 25-38.
- Candel, I., Memon, A., Al-Harazi, F. (2007). Peer discussion affects children's memory reports. *Applied Cognitive Psychology, 21*, 1191-1199.
- Ceci, S.J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403-439.
- Cohen, R.L., & Harnick, M.A. (1980). The susceptibility of child witnesses to suggestion. *Law and Human Behavior, 4*, 201-210.
- Cutler, B.L., Penrod, S.D., & Stuve, T.L. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behaviour, 12*, 41-55.

- Dickinson, J.J., Poole, D.A., & Laimon, R.L. (2005). Children's Recall and Testimony. In N. Brewer & K. D. Williams (Eds.), *Psychology and Law: An Empirical Perspective* (pp. 151-177). New York: Guilford Press.
- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology, 17*, 533-543.
- Gee, S., & Pipe, M.-E. (1995). Helping children to remember: The influence of objects cues on children's accounts of a real event. *Developmental Psychology, 31*, 746-758.
- Goodman, G.S., & Melinder, A. (2007). Child witness research and forensic interviews of young children: A review. *Legal and Criminological Psychology, 12*, 1-19.
- Greenhoot, A.F. (2000). Remembering and understanding: The effects of changes in underlying knowledge on children's recollections. *Child Development, 71*, 1109-1328.
- Hafstad, G.S., Memon, A., & Logie, R. (2004). Post-identification feedback, confidence and recollections of witnessing conditions in child witnesses. *Applied Cognitive Psychology, 18*, 901-912.
- Hoffman, H.G., Granhag, P.A., Kwong See, S.T. & Loftus, E.F. (2001). Social influences on reality-monitoring decisions. *Memory & Cognition, 29*, 394-404.
- Holliday, R.E. (2003). Reducing misinformation effects in children with cognitive interviews: Dissociating, recollection and familiarity. *Child Development, 74*, 728-751.
- Hope, L., Ost, J., Gabbert, F., Healey, S. & Lenton, E. (2008). "With a little help from my friends...": The role of co-witness relationship in susceptibility to misinformation. *Acta Psychologica, 127*, 476-484.
- Jonsson, A.C., & Allwood, C.M. (2003). Stability and variability in the realism of confidence judgments over time, content domain and gender. *Personality and Individual Differences, 34*, 559-574.

- Juslin, P., Olsson, H., & Winman, A. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review*, *107*, 384-396.
- Lamb, M.E., Sternberg, K.J., Orbach, Y., Esplin, P.W., Stewart, H., & Mitchell, S. (2003). Age differences in young children's responses to open-ended invitations in the course of forensic interviews *Journal of Consulting and Clinical Psychology*, *71*, 926-934.
- Lichtenstein, S., Fischhoff, B., & Phillips, L.D. (1982). Calibration of probabilities: The state of the art 1980. In P.S. Kahneman & A. Tversky (Eds.), *Judgments under uncertainty: heuristics and biases* (pp. 306-334). New York: Cambridge University Press.
- Lindberg, M.A., Chapman, M.T., Samscock, D., Thomas., S.W., & Lindberg, A.W. (2003). Comparisons of three different investigative interview techniques with young children. *Journal of Genetic Psychology*, *164*, 5-28.
- Lindsay, R.C.L., Wells, G.L., & Rumpel, C.M. (1981). Can people detect eyewitness identification accuracy within and across situations? *Journal of Applied Psychology*, *66*, 79-89.
- Loftus, E F., & Hoffman, H.G. (1989). Misinformation and memory: the creation of new memories. *Journal of Experimental Psychology: General*, *118*, 100-104.
- McClelland, A.G.R., & Bolger, F. (1994). The calibration of subjective probabilities: Theories and models 1980-1993. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 453-481). New York: John Wiley & Sons.
- Meade, M.L., & Roediger, H.L., III. (2002). Explorations in the social contagion of memory. *Memory & Cognition*, *30*, 995-1009.
- Olsson, N., & Juslin, P. (2002). Calibration of confidence among eyewitnesses and earwitnesses. In P. Chambers & M. Izaute (Eds.), *Metacognition: Process, function and use* (pp. 203-218). Dordrecht, Netherlands: Kluwer Academic Publishers.

- Pansky, A., Koriat, A., & Goldsmith, M. (2005). Eyewitness recall and testimony. In N. Brewer & K. D. Williams (Eds.), *Psychology and Law: An Empirical Perspective* (pp. 93-150). New York: Guilford Press.
- Paterson, H.M., & Kemp, R.I. (2006). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime & Law, 12*, 181-191.
- Penrod, S.D., & Cutler, B.L. (1995). Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy and Law, 1*, 817-845.
- Pipe, M.-E., Lamb, M.E., Orbach, Y., & Esplin, P.W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review, 24*, 440-468.
- Poole, D.A., & Lindsay, R.C.L. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology, 7*, 27-50.
- Roebbers, C.M., Gelhaar, T., & Schneider, W. (2004). "It's magic!" The effects of presentation modality on children's event recall: Developmental progression in the impact of question format. *Journal of Experimental Child Psychology, 87*, 320-335.
- Roebbers, C.M., & Howie, P. (2003). Confidence judgments in event recall: Developmental progression in the impact of questions format. *Journal of Experimental Child Psychology, 85*, 352-371.
- Rudy, L., & Goodman, G.S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology, 27*, 527-538.
- Sternberg, K.J., Lamb, M.E., Orbach, Y., Esplin, P.W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of Applied Psychology, 86*, 997-1005.

- Sutherland, R., Pipe, M.-E., Schick, K., Murray, J., & Gobbo, C. (2003). Knowing in advance the impact of prior event information on memory and event knowledge. *Journal of Experimental Psychology*, *84*, 244-263.
- Wells, G.L., Lindsay, R.C.L., & Ferguson, T.J. (1979). Accuracy, confidence, and juror perception in eyewitness identification. *Journal of Applied Psychology*, *64*, 440-448.
- Yates, J.F. (1994). Subjective probability accuracy analysis. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 381-410). New York: John Wiley & Sons.

Appendix A



Kära föräldrar

Vi heter Sandra MacLeod och Sandra Buratti och vi går sista terminen på psykologprogrammet vid institutionen för psykologi, Lunds universitet. För tillfället arbetar vi med vårt examensarbete i vittnespsykologi. Vi studerar alltså vid Lunds Universitet och vårt projekt handleds av professor Carl-Martin Allwood. Vårt arbete består av en undersökning som studerar barns minnesbilder av olika händelser och hur pass säkra de är på sina minnesbilder. Det är viktigt att genomföra följande försök för att få kunskap som säkerställer barn förmåga att vittna. Trovärdigheten för barns vittnesmål ökar med mer kunskap.

Denna studie går ut på att barn i åldern 9-10, två och två, skall få titta på en film som föreställer en picknick där olika vardagliga händelser utspelar sig. I filmen förekommer **inget våld** eller omoraliska scener. Efter barnen har sett filmen träffar vi barnen ett och ett. De skall då fritt berätta vad de kommer ihåg av filmen. Detta spelas in på bandspelare. Efter det ber vi barnen fylla i en enkät med frågor om vad det har sett på filmen, samt hur säkra de är på sina minnen. Efter en vecka kommer vi tillbaka och ber barnen fylla i ännu en enkät. Sammanlagt rör det sig om två tillfällen.

Barnens anonymitet kommer att skyddas helt. Under själva försöket har endast undertecknade kännedom om deras identitet och deras svar märks med en kod. När försöken är klara och vi ska analysera materialet slängs manualen för kodningen och inte ens vi vet vilket barn som står för vilken kod. ***Det är viktigt att ni inte beskriver något om försöket för Era barn, eftersom de skall vara oförberedda inför filmen.*** Klassföreståndaren har redan gett nödvändig information till barnen.

För att Ert barn skall kunna medverka behöver vi Er och Ert barns påskrifter på lappen som medföljer. ***Inget barn deltar utan att vi har fått tillstånd av föräldrarna samt att barnen själv är villiga att delta.*** Lappen lämnas därefter till klassföreståndaren för vidare befordran till oss. Har du några frågor kan du kontakta oss på telefon 0739-380884 (Sandra M) eller 073-2521215 (Sandra B) eller via mail: vittnespsykologi@yahoo.se

Med varma hälsningar

Sandra Buratti

Psykologkandidat

Email: sandra.buratti.516@student.lu.se

Sandra MacLeod

Psykologkandidat

Email: sandra.macleod.590@student.lu.se

Handledare:

Carl Martin Allwood

Professor i kognitiv psykologi

Email: cma@psychology.lu.se

Barnets namn:

Jag vill delta i försöket

Jag vill **inte** delta i försöket

.....
Barnets underskrift

Det är okej att mitt barn deltar

Det är **inte** okej att mitt barn deltar

.....
Målsmans underskrift

Appendix B

1. Discuss and tell each other what the different characters in the film did!
2. Review the sentences below! Discuss and tell each other which of these events took place in the film. You do not have to agree on the answers.
 - Somebody read a book?
 - Somebody took a photograph?
 - Somebody played cards?
 - Somebody received a present?
 - Somebody was knitting?
 - Somebody received a phonecall?
 - Somebody was armwrestling with someone?
 - Somebody got a hug?
 - Somebody was eating icecream?
 - Somebody kicked a ball?

Appendix C

- A. Discuss together and tell each other about your favorite subject in school.
- B. Discuss and tell each other: what is the best thing about your favorite subject?
- C. Make a list together of 5 fun things you like to do after school.

1. _____

2. _____

3. _____

4. _____

5. _____