

# Children's False Memory for Emotional Events

A Developmental Perspective on Emotion's Impact on  
Backwards Causal-Inference Errors

Kristin Kragh

&

Frida F. F. Vennerød



Master of Philosophy in Psychology,

Developmental Psychology

Department of Psychology

UNIVERSITY OF OSLO

May 2014

## Acknowledgements

First and foremost, we would like to thank our supervisors, post doctor Else-Marie Augusti and Professor Annika Melinder, at the Cognitive Developmental Research Unit (EKUP), Department of Psychology, University of Oslo. Thank you, Annika, for using your vast knowledge and experience to help us conduct the project. In all phases, from planning, recruitment, testing to analyzing, your help has heightened the quality of our research. Else-Marie, we want to thank you for your positive spirit, close supervision, and always stressing that less is more. We are also grateful for interesting discussions, particularly regarding merits and disadvantages of different statistical measurements. For mental support and helpful guidance, we want to thank Research Fellow Miriam Sinkerud.

The present study is part of a larger collaboration between EKUP and the University of Padova that we are very thankful to have participated in. As the present study is the first use of the paradigm with Norwegian school children, we wish to thank the two primary schools for participating. A special thank you to head masters Ida Nordstrand and Ola Øvergaard for believing in the importance of the present research, as well as your cooperation and hospitality during the four months of testing. We also want to pay regards to the 132 children we have tested – thank you for being helpful and observant.

Johannes, you have been our mascot throughout the entire study. We want to thank you for helping us understanding the experiment from a child's perspective. Special thanks to Øyvind, for sharing his deep insight regarding numbers, and reviewing the thesis.

Kristin: First of all, I would like to thank my dear partner Petter for always being supportive, encouraging and patient with me during the last two years, as well as several interesting discussions regarding memory. I would also like to thank my beloved children Johannes and Kristian for inspiring me to always do my best, and never give up. Last, but not least, I would like to thank my mother Guðrún for being a good listener, showing much interest and support of my work.

Frida: To my parents Jorunn and Christian, sister Vivian, brother Ø, and partner Simon, for being both supportive and argumentative, and thus challenging me in our discussions about absolutely everything, particularly in those about psychology and memory. I can never thank you enough. V & Ø – you help me better remember and understand the verbatim and the gist. Mum and Dad, everything I know and can do, is because of you. Lastly, I would like to thank S: your words can induce super powers.

## Abstract

The present study examines how emotion affects false memory formation using the backwards causal-inference paradigm, with a developmental perspective. One-hundred-and-thirty-two children participated in the study, with 56 children aged 6-8 years, 43 children aged 9-10 years and 33 children aged 11-12 years. The children were presented with one of six different PowerPoints, which all displayed the same scripts in photographs, but differed in emotional (positive vs. negative vs. neutral) outcome of the script. After a retention interval, the children were presented with 30 photographs previously displayed, and 30 photographs that were new. The children then expressed which of the pictures they had seen or not with a simple “yes” or “no” response. Children aged 9-10 were significantly better at discriminating between old and new pictures compared to children aged 6-8. The children aged 11-12 were not significantly different from the other age groups. No significant developmental reversals were revealed. Regardless of age, the children made significantly fewer gap-filling errors than causal errors. There was a significantly lower rate of memory errors for positive stimuli than for both negative and neutral stimuli, and a significantly more liberal bias for positive and negative scripts than for neutral scripts when using the response criterion. No significant main effect of emotion was found using the discrimination accuracy index.

## Table of Contents

<b>Introduction</b> .....	<b>1</b>
<b>Human Memory and its General Development</b> .....	<b>1</b>
Fuzzy-Trace Theory.....	3
Developmental Reversals.....	4
The Deese Roediger-McDermott-Paradigm.....	4
Two Different Mechanisms: Gap-Filling Errors and Causal Errors.....	5
<b>The Backward Causal-Inference Paradigm</b> .....	<b>6</b>
<b>How does Emotional Memory Differ from Neutral Memory?</b> .....	<b>7</b>
Positive vs. Negative Memories.....	8
Valence and Arousal.....	8
Why is there an Effect of Emotionality on False Memory?.....	9
<b>The Present Study</b> .....	<b>11</b>
<b>Method</b> .....	<b>13</b>
<b>Participants</b> .....	<b>13</b>
<b>Materials</b> .....	<b>13</b>
Pictorial Stimuli and Apparatus.....	13
Recognition Phase.....	14
<b>Procedure</b> .....	<b>15</b>
Encoding Phase.....	15
Retention Interval.....	16
Recognition Phase.....	16
<b>Ethical Considerations</b> .....	<b>16</b>
<b>Statistical Analyses</b> .....	<b>17</b>
Proportion Scores.....	17
Discrimination Accuracy Index and Response Criterion.....	17
<b>Pilot Testing</b> .....	<b>19</b>
<b>Results</b> .....	<b>22</b>
<b>Estimating Sample Size</b> .....	<b>22</b>
<b>Preliminary Analyses</b> .....	<b>22</b>
<b>Main Analyses</b> .....	<b>23</b>
Proportion Scores.....	23
Discrimination Accuracy Index.....	25
Response Criterion.....	27
<b>Discussion</b> .....	<b>28</b>
<b>Age Related Performance in Recognition Memory</b> .....	<b>28</b>
<b>False Memories and Age</b> .....	<b>29</b>
<b>Divergent Influences of Emotionality on Memory</b> .....	<b>32</b>
<b>Limitations of the Present Study</b> .....	<b>35</b>
<b>Future Directions</b> .....	<b>36</b>
<b>Concluding Remarks</b> .....	<b>37</b>
References.....	<b>38</b>
Appendix.....	<b>46</b>

## **Introduction**

Human memory is generally well adapted to everyday situations through evolution (Schacter, 1996), but even so, it is not without faults. During the last decades, the nature of memory has been questioned in several court trials, particularly in those involving children as witnesses. In infamous and controversial cases such as the Bjugn-affair (see Magnussen, 2004), the Little Rascals preschool case, and the McMartin Preschool Trial (for reviews, see Sabbagh, 2009; Schacter, 1996) the suggestibility and credibility of children has been scrutinised. This has generated extensive research and debate about the formation of false memory, a dispute so controversial that it has been called “The Memory Wars” (Schacter, 1996). Even so, there is still not consensus on how emotions affect memory errors from a developmental perspective (Brainerd & Reyna, 2005; Brainerd & Reyna, 2012). In a courtroom, the common denominator is often the emotionality of the event investigated. How development affects the impact of emotion on memory errors is therefore of utter importance both in forensic situations and from a scientific perspective (Brainerd, Holliday, Reyna, Yang, & Toglia, 2010).

Memory is part of a cognitive system, capable of both constructing and reconstructing information (Brainerd & Reyna, 2005). Similar processes thus generate both accurate and inaccurate memory. The present study takes place in a controlled environment, which makes it readily to evaluate whether reports of scripted events are accurate or not. This is opposed to reports in a courtroom where there is no certain way of knowing if a statement is affected by memory distortions. The paradigm employed investigates inferential memory errors via the incorrect retrieval of a presented story-script, measuring how emotions affect memory for causal inferences and gap-filling errors in children.

### **Human Memory and its General Development**

Episodic and semantic memories are both forms of explicit or declarative memory. Episodic memory is defined as the retrieval of events and previous experiences, whereas semantic memory is our generic, factual knowledge about the world (Goswami, 2008). These memories can be brought to mind consciously and deliberately. Explicit memories are often contrasted with implicit memories, which are procedural memories that make changes in performance without the involvement of consciousness (Schacter, 1996). In cognitive psychology, memory is assumed to be a modular system, with long-term memory divided

into explicit and implicit memory, and episodic-, semantic- as well as procedural memory considered to be distinctive in various ways. Thus, these different types of memory are usually considered independently of each other in cognitive research. Further, this modular approach is supported by cognitive neuroscience, with different types of memory appearing to rely on different anatomical structures in the brain (Goswami, 2008). The present study aims at addressing a phenomenon pertaining to episodic memory.

Children and adults construct their episodic memories, and that process depends on prior knowledge and personal interpretations (Goswami, 2008). From early on, children remember events that are important in their everyday life, merging these memories into scripts such as the usual bedtime-routine, and episodic memory thus develops (Nelson, De Haan, & Thomas, 2006). The construction of memories may also depend on how much sense a person can make of their experiences (Schacter, 1996). Very young children may not structure their experiences in memorable ways, especially if they do not understand their own experiences or do not have a clear temporal framework. Memory development is integrated in larger social and cognitive activities, and should therefore be looked at in context with other developmental processes.

Several studies have used tasks and procedures derived from larger social and cognitive developmental perspectives, attempting to provide a purer measure of a particular memory system of interest in order to figure out its contribution to human memory (for a review, see Brainerd & Reyna, 2012). Episodic memory is typically measured by standard tests of recall and recognition (Cooper & Schacter, 1992; Gallo, 2013; Nelson et al., 2006). Recall memory is the ability to actively recall previous experiences (Goswami, 2008), whereas recognition memory is the ability to identify a stimulus as previously encountered (Ghetti & Angelini, 2008).

Until recently, most attention has been allocated to research of true memory (Brainerd & Reyna, 2005). However, cases of people recovering previously unsuspected memories of childhood abuse, sketchy interrogations, confessions based on unreliable reports, and leading questionings, especially with children involved, has turned the attention towards false memories (Brainerd & Reyna, 2005; Ceci & Bruck, 1995; Magnussen, 2004). False memory refers in a very general way to circumstances where one is convinced of definitive memories of an event that did not actually happen (Brainerd & Reyna, 2005). Fuzzy-trace theory (FTT) is a contemporary theory developed by Brainerd and Reyna (1995) that explains false-memory phenomena.

## **Fuzzy-Trace Theory**

FTT posits two types of memory processes, and has therefore been referred to as a dual process theory (Brainerd & Reyna, 2002). The core idea of FTT is that false-memory responses are affected by two mechanisms; *recollection* and *familiarity*, which operate in opposition to each other (Brainerd & Reyna, 2005). Recollection involves recovery of qualitative information associated with the context in which the event was encountered. As usually prompted by a critical cue, this induces a subjective sense of vivid reliving of one's personal past (Brainerd, Holliday, & Reyna, 2004). A familiarity process, on the other hand, allows for a general sense of "oldness" about an item, where a memory trace is assessed without the ability to retrieve any qualitative details about the specific event (Lyons, Ghetti, & Cornoldi, 2010).

A variety of evidence concludes that the human brain deposits dissociated *verbatim* and *gist* traces of experience (Reyna & Brainerd, 1995). FTT supposes that verbatim traces are detailed integrated representations of several surface features, and other item-specific information, requiring a recollective retrieval (Brainerd & Reyna, 2004; Brainerd, Reyna, & Forrest, 2002). The retrieval of such traces in memory tests induces a vivid mental reinstatement of experiences accompanied by targets of earlier presentations. Forgetting of verbatim traces thus creates disintegration of features (Reyna & Titcomb, 1997). On the other hand, the representations of gist traces are episodic interpretations of concepts as meanings, relations, and patterns (Brainerd & Reyna, 2005). These concepts are elaborations generated by a person while encoding the targets' surface form.

The main difference between verbatim and gist information is that verbatim information is part of the memory experienced by the subjects (i.e., remembering details of the story and what the photographs looked like), whereas gist information must be added to the experience by the subjects themselves (i.e., the general meaning of the script) (Brainerd & Reyna, 2004). The initial encoding of experiences' surface features (verbatim trace) is assumed to initiate a corresponding mechanism of meaning access and elaboration (gist trace), and these different mechanisms of storage run in parallel while encoding continues.

FTT assumes that in true recall or recognition of verbatim traces, a lifelike form of remembering occurs, where subjects feel they are consciously re-experiencing targets' occurrence in specific contexts (Brainerd & Reyna, 2005). This is traditionally called recollection in dual process theories. Verbatim retrieval is assumed to predominate with targets, but sometimes the targets may fail to evoke verbatim retrieval and may instead evoke

gist retrieval. Gist retrieval usually evokes a more global and inchoate form of remembering, called familiarity in dual process theories. In false memory, verbatim and gist retrieval are assumed to be opponent processes, with gist retrieval supporting false-memory responses, and verbatim retrieval suppressing them (Brainerd & Reyna, 2005).

### **Developmental Reversals**

Cognitive factors as source-monitoring, social factors as susceptibility to persuasion, metacognitive factors as introspective awareness of memory state, and neurobiological factors as maturation of the prefrontal cortex, all contribute to promote memory accuracy and resistance to suggestion (Brainerd, Reyna, & Ceci, 2008). Together these processes ensure age declines in memory distortions. Several studies have found robust increases in recognition and recall memory, in part due to the development of the above-mentioned factors (Ackerman, 1984; Davidson & Hoe, 1993; Lindsay, Johnson, & Kwon, 1991).

In recent years, exceptions have been found to the general allegation of age decline in false memory. These exceptions have been predicted by the FTT (Brainerd & Reyna, 1998). Several studies have found methods that detect age increases in false memory, and some even with a net decline in accuracy (see Brainerd, Reyna, & Ceci, 2008, for a review; Lyons et al., 2010). Previous research indicates that the propensity to form false memories decreases with age when children are exposed to social pressure (Bruck & Ceci, 1999) or when they fail to correctly recognize the source the information originally was coming from, especially when the sources are highly similar to one another (Lindsay et al., 1991). However, when children falsely remember the underlying gist of events, developmental increases in false memories have been observed (Brainerd & Reyna, 2012; Brainerd, Reyna, & Ceci, 2008; Lyons et al., 2010). Several studies supporting the FTT have found age increases in false memory to be due to improvement in children's memory for semantic gist (Brainerd et al., 2010; Gallo, 2013). Because of the increase in false memory formation as a function of age, this developmental pattern has acquired the name "developmental reversals" (Brainerd, Reyna, & Ceci, 2000).

### **The Deese Roediger-McDermott-Paradigm**

*The Deese Roediger-McDermott paradigm* (DRM), created by Deese (1959) and Roediger and McDermott (1995), is a frequently used paradigm that has shed some light on the development of false memory formation, supporting the predictions of FTT (Anastasi & Rhodes, 2008; Brainerd, Forrest, Karibian, & Reyna, 2006; Dewhurst, Pursglove, & Lewis,



2007; Howe, Cicchetti, Toth, & Cerrito, 2004) This paradigm is carried out by the use of word lists where a familiar stimulus word like “doctor” is chosen, and a word list is created by selecting 12-15 semantic associates to that word (nurse, needle, hospital, sick etc.) These associated words are then presented to the participants without the stimulus word “doctor” (i.e., the critical lure). Subsequently after having studied the word list, the participants make old/new recognition judgments of actually studied words, critical lures, and other lures that are not semantically related (e.g., ocean) to the studied materials. Older children and adults have been shown to falsely recognize critical lures and correctly recognize studied words (Roediger & McDermott, 1995). This is due to older children’s and adult’s ability to connect the meaning of studied words, leading the critical lure to be coded along with the items studied during the encoding phase. In the DRM-paradigm, this effect has been shown to increase with age, causing an age-increase in memory distortion (Brainerd et al., 2004; Brainerd & Reyna, 2012; Brainerd, Reyna, Ceci, & Holliday, 2008; Holliday, Brainerd, & Reyna, 2011; Howe, 2007; Metzger et al., 2008; but see Ghetti, Quin, & Goodman, 2002).

Firm conclusions on developmental reversals can however not be drawn from studies using the DRM-paradigm alone, as age-increases in memory errors using this paradigm may also be due to age-related increases in children’s ability to detect relations among items on the word-list (e.g. Brainerd, Reyna, & Ceci, 2008; Howe, 2007; Lyons et al., 2010). Thus, a new memory paradigm has been developed to better study the developmental reversals suggested in FTT (Hannigan & Reinitz, 2001; Lyons et al., 2010).

### **Two Different Mechanisms: Gap-Filling Errors and Causal Errors**

Hannigan and Reinitz (2001) found evidence of two different mechanisms underlying inference-based memory errors, with schema-based memory errors (*gap-filling errors*) likely to be fundamentally different from errors based on causal inference (*causal errors*). Gap-filling errors can be defined as erroneously remembering script-consistent distractor images, whereas causal error is erroneously remembering the not seen cause of previously seen effects (Mirandola, Toffalini, Grassano, Cornoldi, & Melinder, 2013).

Hannigan and Reinitz (2001) found that gap-filling errors were associated with “know” responses, indicating that they tended to be based on a feeling of familiarity. Causal errors were however associated with “remember” responses, indicating that they tended to be based on an explicit recollection from the encoding phase. This led the authors to the conclusion that recollection supports causal-errors, and that familiarity supports gap-filling errors. Hannigan and Reinitz (2001) thus suggest that gap-filling errors resulting from

familiarity is due to an activation of generic semantic information or knowledge in memory. Further, they theorize causal errors to occur when people falsely recall having experienced a specific item, and thus not remembering that the memory is actually an inference they made themselves based on a picture which showed the effect of an event during encoding

### **The Backward Causal-Inference Paradigm**

Based on the above-mentioned dissociations in recognition memory, a relatively new paradigm named *the Backward Causal-Inference Paradigm* was developed by Hannigan & Reinitz (2001). This paradigm has later been applied to further explore the effects of age-related increases in false-memory formation (Lyons et al., 2010). In this paradigm, series of photographs are composed together to form a story-script (e.g. eating at a restaurant). The series include photographs of effects (e.g., wiping up water from a table) and non-presented causes (e.g., knocking over a glass of water), as well as script-consistent and script-inconsistent photographs. After a retention interval the participants complete a yes/no recognition test. The test includes the causes (e.g. knocking over a glass of water) whose effect-photograph had been seen during encoding (e.g. wiping up water from a table), as well as new and old script-consistent and script-inconsistent photographs (Hannigan & Reinitz, 2001; Lyons et al., 2010; Mirandola, Paparella, Re, Ghetti, & Cornoldi, 2012; Mirandola et al., 2013).

Hannigan and Reinitz (2001) originally presented the material to adults testing for memory-distortion, but the backward causal-inference paradigm also allows testing for age differences. This was first done by Lyons and colleagues (2010), who presented the paradigm to typically developing children. In contrast to the DRM-paradigm, the backward causal-inference paradigm does not rely on the ability to detect the overall theme of a word-list across multiple items, and may therefore provide a better distinction between the developments of recollection- and familiarity-based errors (Lyons et al., 2010).

With the backwards-causal inference paradigm being relatively new, few studies have been conducted using this paradigm on adults, and even fewer with the aim of exploring the development of memory distortions in children. The former mentioned study of Lyons and colleagues (2010) tested 6-, 7-, 9-, 10-, and 18-year olds for memory distortions focusing on backward causal errors and gap-filling errors. As predicted, they found that although age-related increases in backward causal inference errors were observed, gap-filling errors were age-invariant. Six year olds did not show backward causal inference errors, whereas the older

groups did. Adults were significantly more likely to show this effect compared to the youngest children. A significant age-increase in causal errors between six year olds and nine year olds and adults was detected, and all age groups except six year olds showed backward causal-inference errors significantly different from zero. Regarding gap-filling errors, no significant age-related differences were detected.

In a study by Mirandola and colleagues (2012) the paradigm was presented to children with attention deficit/hyperactivity disorder. Later, a modified version of the paradigm adapted to include a difference in emotional content was presented to adults (Mirandola et al. 2013). Another study using the Backwards Causal-Inference Paradigm with emotional content included children with learning disabilities, but did not include a developmental perspective (Mirandola, Losito, Ghetti, & Cornoldi, 2014).

Developmental improvements in the ability to form meaning connections between experiences are necessary to produce age-increases in false memory (Brainerd, Reyna, & Ceci, 2008). The enhancement in forming of meaningful connections is predicted to lead to a developmental increase in backwards-causal memory errors. This is because older children may falsely recollect the causal photographs, confusing them for causal inferences they made when viewing corresponding effect photographs during encoding (Lyons et al., 2010).

### **How does Emotional Memory Differ from Neutral Memory?**

Emotionality has been indicated to influence explicit memory regarding the number of events remembered as well as the quality and accuracy of memory (for a review, see Kensinger & Schacter, 2008). Emotional events have repeatedly been found to be better remembered than neutral events (for reviews, see Brainerd & Reyna, 2012; Buchanan & Adolphs, 2002) in both children and adults, and in the laboratory as well as in more naturalistic settings (Burgwyn-Bailes, Baker-Ward, Gordon, & Ornstein, 2001). The heightened true memory for emotional stimuli has been suggested attributed to the higher evolutionary benefit of recognizing and remembering potentially dangerous or rewarding situations rather than neutral situations (Wright, 1994). Regarding false memory, differences in memory errors have been suggested to be an effect of more fluently processing negative stimuli, leading people to falsely believe that they recognize an item (Kensinger & Schacter, 2008). Research on the effect of emotionality on memory has mostly focused on negative emotion, and more precisely, on eyewitness-memory, flashbulb memories, and traumatic memories (Schooler & Eich, 2000).

### **Positive vs. Negative Memories**

Flashbulb-memory (Brown & Kulik, 1977) was an early concept in the field of emotional memory research. These memories are described as highly emotional, and the theory states that the traumatic content enhances how much one might remember. However, later research has suggested strong source amnesia for these events. One might believe that the memory is very strong and robust for one's own flashbulb-memories, but the subjective confidence in this is often not matched by the objective accuracy (Schacter, 1996). This indicates how the strong emotionality of a memory might intervene with accuracy, making it a false memory. Research with children has suggested traumatic memory to be influenced by the same variables as non-traumatic memory, but that the distinctiveness of the situation will enhance the memory (for a review, see Cordon, Pipe, Sayfan, Melinder, & Goodman, 2004).

As the majority of research on emotional memories investigate differences between neutral and negative memories, less is known about the effect of positive emotion on memory. There has however been a tendency to focus more on the effects of positive emotion in recent years (Brainerd et al., 2010; Brainerd, Stein, et al., 2008). It is important to differentiate not only between neutral and emotional stimuli, but also between positive and negative, as these emotions have in the limited research conducted been indicated to have opposite effects on familiarity and use of verbatim suppression (Brainerd et al., 2010; Brainerd, Stein, et al., 2008). In a previous study on children's episodic memory, children reported their own narratives for either positive or negative events (Fivush, Hazzard, McDermott Sales, Sarfati, & Brown, 2003). The children reported more objects, people and details for positive experiences, and more information of their own thoughts and emotions for negative events. Controlled research on memory differences for positive and negative story-scripts could therefore help to further understand differences in how children both accurately and falsely remember events that differ in emotionality.

### **Valence and Arousal**

Emotionality does not only differ between positive or negative, but can also be divided into *valence* (emotional response) and *arousal* (intensity of response). Both the valence of stimuli and the level of arousal it induces has been shown to influence memory (Bradley & Lang, 1994), but with divergent effects (Kensinger, 2004). The general tendency is towards negatively valenced pictures being remembered better than neutrally valenced pictures (Christianson, 1992; Cordon, Melinder, Goodman, & Edelstein, 2013; Kensinger &

Corkin, 2003; Oschner, 2000), and high and moderate arousing stimuli remembered better than low arousing stimuli (Brainerd et al., 2010; Cordon et al., 2013; Kensinger & Corkin, 2003; Oschner, 2000) for both children and adults.

Valence and arousal are variables used in several studies, but a minority of research do control for the different influences they have on memory (Kensinger, 2004). In such controlled studies with adults, the effect of arousal has been indicated to be larger than the effect of valence (Kensinger, 2004; Kensinger & Corkin, 2003). The sparse age-normed valence and arousal research in children has indicated less clear tendencies (Brainerd et al., 2010; Cordon et al., 2013). Age-norming emotional stimuli is important as valence and arousal ratings seem to be a function of age, and the validity of studies without age-normed material have thus been questioned (Cordon et al., 2013).

### **Why is there an Effect of Emotionality on False Memory?**

The effect of differently valenced stimuli on false memory is proposed explained by FTT (Brainerd et al., 2010; Brainerd, Stein et al., 2008). Brainerd, Stein and colleagues (2008) used recognition of DRM-lists with different valences that did not differ regarding arousal. The results indicated a higher degree of false recognition for negatively valenced compared to neutrally valenced distractors, which is the common tendency (see Kensinger & Schacter, 2008, for a review). They also found more memory errors for neutral than for positive distractors. In line with FTT, the authors accounted this to be a factor of the impact of valence on both *similarity judgment*, which is falsely accepting seemingly familiar material by processing gist traces, and *recollection rejection*, which is the tendency to use verbatim-memory to suppress a false acceptance of critical distractors. The similarity parameter for critical distractors increased from positive to negative distractors, with neutral distractors in the middle. Therefore, negative valenced stimuli induced the highest level of familiarity in the participants, with positive stimuli inducing the lowest level of familiarity. The level of recollection rejection decreased from positive to negative valence, with again, neutral scoring in between. Therefore, negatively valenced stimuli shows the highest level of false recognition as it both induces a high level of meaning familiarity, and less ability to use verbatim memory to suppress the false acceptance of these seemingly familiar items (Brainerd, Stein et al., 2008)

As this different effect of emotional stimuli on memory errors was found in university students, (Brainerd, Stein, et al., 2008) the development of this effect has been further investigated. As an effect of FTT, this tendency is thought to increase with age, with a

positive correlation between development and false memory for negative stimuli (Brainerd & Reyna, 2012). According to FTT, emotional valence will result in the same developmental trends as semantic gist will, as it is a conceptual aspect of the event (Brainerd & Reyna, 2012; Brainerd, Stein, et al., 2008; Rivers, Reyna, & Mills, 2008). Therefore, the valence of an emotional situation will lead to a developmental increase in false memories as a factor of gist, as the ability to catch the meaning will increase with age. Arousal is viewed as a concept of verbatim memory, as it is non-conceptual (Brainerd & Reyna, 2012). Higher arousal in an emotional situation will also lead to a developmental increase in false memories, but as a factor of verbatim experience.

Supporting this, an age-increase in false memory from the age of seven to the age of 11 years, and onwards to young adults, has been found for negative stimuli (Brainerd et al, 2010). The higher degree of false memories for negative stimuli relatively to neutral or positive stimuli is hypothesized to emerge between the ages of seven and 11 years. This is supported by no significant difference in memory errors for positive and negative stimuli for the seven-year-olds, and 11-year-olds having a significantly higher degree of false memories for negative than for positive stimuli (Brainerd et al, 2010). Between the age of 11 and 20 years, false memories increased the same amount for both negative and positive stimuli, indicating that the differentiation in memory errors between negative and positive stimuli occurs before the age of 11 years.

Fitting with this theory of a developmental increase in emotional false memories, Howe (2007) found the age-increase in false recognition memory to be highest for negative rather than for neutral stimuli. The tendency has been, as suggested by several other recognition studies, for children to have an increase in false memories with age, especially for negative stimuli (see Brainerd & Reyna, 2012, for a review; Howe et al., 2010; Mirandola et al., 2014). However, the tendency for a lower proportion of inferential errors in negative stimuli rather than neutral stimuli has been reported in other studies with children (Ceci, Lofthus, Leichtman, & Bruck, 1994; Mirandola et al., 2013; Otgaar, Peters, & Howe, 2012). Mirandola and colleagues (2013) attributed this finding to a finer discrimination for emotional stimuli than for neutral stimuli.

Conflicting data has thus been found on the effect of emotion on discriminability. Reduced discriminability for emotional stimuli has been found, with children having a reduced capacity to discriminate false from true memories in emotional events (Howe, 2007), and a higher capacity for discriminating in neutral events. A more liberal bias for negative stimuli than for neutral stimuli has also been found (Mirandola et al, 2013), indicating that

children are more prone to falsely accept negative stimuli. These divergent results might be a factor of different types of stimuli used in testing.

In developmental research, stronger effects of valence have been found with recall memory than with recognition memory (Burell, 2013; Ceci & Bruck, 1993, but see Howe, 2007). Recognition memory might not be as influenced by valence because it is possible to truly recognize stimuli without correctly remembering the source of the memory (Kensinger & Schacter, 2008), and thus not remember the emotion. It is nevertheless important to keep in mind that the Backwards Causal-Inference Paradigm using pictures of varying valence has at the present date not been used in developmental research before, so the conclusions for the above mentioned studies might not apply for the present study.

## **The Present Study**

The aim of this study is to examine inferential memory error (i.e., gap-filling errors and causal errors) in recognition memory of a presented picture story-script. Children aged six to 12 years were presented with six different story-scripts of positive, negative, and neutral content. After a five-minute retention interval a recognition test was presented where the children discriminated between unseen (new) and previously seen (old) pictures.

Positive content has for the first time been included in this paradigm, and is likely the first of its kind to be used on typically developing children with the aim of examining effects of age-related increases in false-memory formation. The positive content has been added to further investigate the role of emotionality, and to investigate if there is a general difference in impact on false memory between emotional and neutral stimuli, or if there is an added difference between types of emotionality.

The age-range for this study was selected based on theoretical hypotheses suggesting that recollection emerges later, and shows more developmental change than familiarity (Ghetti & Angelini, 2008). Studies have shown that the subjective experience of recollection is not present until a certain age (Perner & Ruffman, 1995). Before children are three to six years old, they may not be able to differentiate among belief, knowledge, and memory, and are therefore not able to experience recollection as a different memory state from familiarity, indicating that recollection emerges from an earlier state of general familiarity (Ghetti & Angelini, 2008).

Based on previous research three main hypotheses were proposed. In the first hypothesis, age-related decreases in overall memory distortions were expected, as numerous

studies show that memory-performance enhances as children develop (Ceci & Bruck, 1993; Lindsay et al., 1991).

The second hypothesis was based on previous research (Lyons et al., 2010) using the same paradigm. There is reason to assume that age-differences should result in age-related differences in the two memory-processes; familiarity and recollection. The familiarity-based memory distortions (gap-filling errors) should be quite age-invariant, whereas recollection-based distortions (causal errors) should increase with age. This age-increase in causal errors could thus mean that younger children will on average have higher proportions of their total correct scores coming from correctly rejecting cause pictures rather than gap-filling pictures.

The third hypothesis states that children, regardless of age, were expected to have more false alarms for negative stimuli, than for neutral stimuli, and the lowest rate of false alarms for positive stimuli. Negative stimuli have been found to induce a higher level of meaning familiarity, and a lesser ability to use verbatim memory to suppress the false acceptance of these seemingly familiar items, therefore causing a higher degree of false alarms (Brainerd, Stein, et al., 2008).



## **Method**

### **Participants**

One hundred and thirty two children participated in the study. Fifty six children were between six and eight years old ( $M = 6.92$ ,  $SD = 0.94$ , 29 males), 43 were between nine and 10 years ( $M = 9.58$ ,  $SD = 0.50$ , 19 males) and 33 children belonged to the group between 11 and 12 years of age ( $M = 11.52$ ,  $SD = 0.51$ , 16 males). One child was eliminated from further analysis, as he was not able to perform the recognition-phase because of a developmental disability. Participants attended one of two elementary schools in two major Norwegian cities, where a total of 250 children were contacted. Both the children and their guardians signed individual consent forms that were delivered to the guardians by envelopes in their children's backpacks.

### **Materials**

Scripts consisting of colour photographs were used. A script is a story shown in pictures, and these scripts depicted people engaged in every day routines, inspired by Hannigan and Reinitz (2001). Originally the paradigm consisted of nine different scripts. A pilot study in two parts was conducted, qualifying six scripts to be part of the present experiment. For further elaboration, see the section of the pilot study.

#### **Pictorial Stimuli and Apparatus**

The apparatus included a computer with a 13'' monitor, set on a table in front of the participant. Each script contained a causal pattern, in which a unique, not shown, cause results in three possible and mutually exclusive consequences; neutral vs. positive vs. negative. For different outcomes in each script, see table 1.

Table 1

*The different outcomes for each script*

Script	Positive picture	Neutral picture	Negative picture
Morning	Female finds roses in a cabinet	Female takes a packet of biscuits out of a cabinet	Female finds a spider in a cabinet
Bicycle trip	Female on a bike meets a friend	Female on a bike entering a building	Female on a bike is injured by a car
Mountain hiking	Male excited to have climbed mountain	Male looking at a leaf	Male falls down from the mountain
Run	Female wins a running competition	Female changing her shoes	Female hurts her ankle, loses the competition
Gambling	Male wins a large amount of money	Male playing on a slot machine	Male loses a large amount of money
Dating	Male and female kiss	Female gives a book to the male	Male hits the female

*Note.* For each script there are three different emotional outcomes.

A series of 21 coloured photographs for each script were created. Eight photographs depicted actions typical for each event, which of one was always shown at the beginning of the script, and one was always shown at the end of the script. Two pictures depicted a positive outcome, two pictures depicted a negative outcome, and two pictures depicted a neutral outcome. One cause picture was created for each script, but not shown during encoding. Additionally, three pictures consistent with the script were labelled as new, and three pictures consistent with the script were labelled as old. The pictures labelled as old and new were used alternatively as target-pictures at encoding or distractors during recognition respectively (See appendix for an example of a script). Finally, stimuli also included 16 photographs inconsistent with any of the scripts (e.g., boy playing chess, people playing on the beach and swimming in the ocean, or simple landscapes). Five of these photographs were shown at the beginning, and five at the end of every presentation during encoding to prevent primacy and recency effects.

### **Recognition Phase**

A unique sequence of 60 photographs in randomized order was shown to all participants. From each script there was:

- a) A representative picture from the script, chosen from the six pictures showing typical series of situations occurring in the script

- b) Three script-consistent photographs from the “new” labelled pictures, serving as a distractor
- c) Three script-consistent photographs chosen from the “old” labelled pictures, serving as a target photograph
- d) One photograph depicting the common cause of the three possible outcomes (never presented at encoding)
- e) One, already viewed, inconsistent photograph serving as an unrelated target.
- f) One new script-inconsistent photograph serving as an unrelated distracter.

The three script-consistent distractor photographs were created for each script by forming two groups of three photographs, alternatively used as a target and distractors, counterbalanced for the conditions. These photographs were randomly presented with the restriction that they would never be placed in either the last or the first position, or to the slide right before the outcome photographs.

## **Procedure**

### **Encoding Phase**

All participants were tested individually at school during school hours. The participants were seated in front of a computer screen that was placed on a table in front of them in a quiet room. They were instructed that they would see a number of photographs displaying people engaging in daily activities, and to watch every photograph carefully to try to understand what the photographs were attempting to convey. The children were not informed that the procedure would measure memory, as incidental encoding will have higher ecological validity than intentional learning. For each script, participants viewed 13 pictures in logical order. Each photograph appeared on the computer screen for two seconds, followed by a black screen also lasting for two seconds. Every participant saw one of the three possible outcomes for each script; two scripts with a positive outcome, two scripts with a neutral outcome, and two scripts with a negative outcome. The overall duration of the encoding phase was approximately six minutes.

### **Retention Interval**

After the encoding phase, the participants were given a filler-task for the duration of a five minutes interval. The filler-task was a puzzle game, consisting of wooden blocks in 3D, and a cardboard where the purpose is to place the blocks within a specific pattern. This filler-task was chosen to distract the children from the scripts they had just seen, and because it was possible to adjust the difficulty level according to age. Whereas the older children were instructed to try to complete the pattern with the blocks, the younger children were instructed to play with the blocks. The atmosphere during the experiment was meant to be comfortable and calm, and therefore small talk was also a part of the retention interval.

### **Recognition Phase**

After the retention interval, the children were asked to resume their position in front of the computer. They were informed that they would see more photographs, some of which they had never seen before, and some of which they had recently seen. For each picture presented, the child was instructed to respond either “yes” or “no” according to the child’s belief in whether the photograph was seen or not during encoding. Two researchers conducted the experiment, of which one always administered the test, whereas the other one always scored the children’s answers. Total duration of testing for each child was approximately 20 minutes.

### **Ethical Considerations**

Before the recruitment of participants started, the study was evaluated by REK (The Regional Committees for Medical and Health Research Ethics), who concluded that the project fell outside the health-research jurisdiction, and did therefore not need any special approval from REK.

Schoolteachers distributed enclosed envelopes, containing a detailed information letter and consent form, for the children to take home to their guardians. The guardians were instructed to sign the consent form if they wished for their child to participate in the study, or refrain from signing if they did not agree to their child’s participation. Instructions requested returning the envelope regardless of their willingness to participate. Guardians were also informed of their possibility to withdraw their children from the project at any time, without having to give any explanation.

Before testing, the children were informed verbally about the project and its purpose. They were also informed about their guardians' consent, and asked to give their own consent to participate by writing down their name or their first letter. All children were informed that they could withdraw from the project at any time during testing without giving any reason.

### **Statistical Analyses**

The children's yes or no responses during the recognition phase were for analysis purposes coded into four different variables.

- 1) Hit – correctly recognizing a previously seen picture
- 2) Correct rejection – correctly rejecting a previously unseen picture
- 3) False alarm – wrongly recognizing a previously unseen picture
- 4) Miss – wrongly rejecting a previously seen picture

### **Proportion Scores**

False recognition was calculated using proportions of raw scores for both false alarm and miss, and is used in the analysis if not otherwise stated.

False recognition was calculated for six conditions respectively; Positive cause pictures, negative cause pictures, neutral cause pictures, positive gap-filling pictures, negative gap-filling pictures, and neutral gap-filling pictures. A 2 (picture type; cause vs. gap-filling) X 3 (emotion; positive vs. negative vs. neutral) X 3 (age; 6-8 vs. 9-10 vs. 11-12 years) repeated measures mixed factorial analysis of variance (ANOVA) with picture-type and emotion as within-group factors, and age as a between-group factor was conducted.

### **Discrimination Accuracy Index and Response Criterion**

The two-high threshold ( $P_r$ ) discrimination accuracy measure is a measure of how precisely participants are able to distinguish old from new pictures in a recognition-test (Snodgrass & Corwin, 1988).  $P_r$  is a common measure of recognition memory, also with children (Augusti & Melinder, 2012; Bayen, Murnane, & Erdfelder, 1996; Johnson, Kounios, & Reeder, 1994; Pollak, Cicchetti, Hornung, & Reed, 2000).  $P_r$  is calculated based on hits and false alarms using the formula [ $P_r = H - FA$ ]. The two-high threshold model does not assume a continuum of memory strength, but defines discrete memory states. In a two-high threshold model there are two memory thresholds; one for old items and one for new items. These two thresholds define three states of memory; old recognition, new recognition, and

uncertainty. A studied item will be accepted if it exceeds the memory threshold, or if it is based on a guess (Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996). Old items crossing the old recognition threshold will therefore always be identified as old, whereas new items crossing the new recognition threshold will always be identified as new, by the participants (Snodgrass & Corwin, 1988). Items in the uncertain state will however be identified as either old or new, depending on the participant's response bias, leading false alarms and misses occurring from an uncertain state.

The response criterion  $B_r$ , [ $B_r = FA/1 - P_r$ ], is therefore defined as the probability of saying "yes" when in an uncertain state of having seen a picture before or not. A value of  $B_r$  equal to 0.5 indicates a neutral bias, a value greater than 0.5 indicates liberal bias in where children are more prone to respond "yes", whereas a value less than 0.5 indicates conservative bias, where children are more likely to respond "no" to whether they have seen a photograph before.

The two-high threshold ( $P_r$ ) discrimination accuracy measure was calculated for the neutral and the emotional scripts (Snodgrass & Corwin, 1988), using corrected hits and false alarms. Similarly the response criterion ( $B_r$ ) was calculated to assess the response bias applied by the children.

## Pilot Testing

As the paradigm is rather new and not extensively used with children, pilot testing was conducted to heighten the reliability of the procedure. This is particularly important in developmental research, as valence and arousal ratings are indicated to be a function of age (Cordon et al., 2013). Because the photographs in the paradigm are developed in Italy, it was necessary to investigate how the procedure would work with Norwegian children, specifically, if the children would comprehend the plot in the scripts as well as the emotions the actors were trying to convey. A two-step pilot testing was performed.

First, a qualitative part, including eight children ( $M$  age = 8.33,  $SD$  = 1.60, 2 girls) was conducted. The children were asked to describe the plot in all nine scripts, as well as explain the emotions of the actors in both the neutral, negative, and positive outcomes. The children were all able to do this, but with two exceptions. Some of the children had problems understanding the dating-script, and some were not able to differentiate the emotions in the neutral and the positive outcomes for both the birthday-script and the boy-script. A difference between Italy and Norway was also found: All the Norwegian children highlighted the fact that the actor in the bicycling-script was not wearing a helmet, and that the actor in the mountain-script did not secure his rope. This indicates that the children were paying close attention to the plots in the scripts. It is also worth mentioning that the gambling-script, which was thought to work less than well in Norway because of the lack of slot machines, did not raise any problems. The children were asked if they understood what the machine was for, and they were all able to explain its function.

The results from the first pilot-test were used as a basis for the second part. The Self Assessment Manikin (SAM), a non-verbal pictorial assessment (Bradley & Lang, 1994; Cordon et al., 2013), was used. With the SAM-assessment manikin, the children ( $N$  = 18,  $M$  = 9.25,  $SD$  = 1.80, 4 boys) rated both picture valence and arousal on a scale from 1-9.

The cut-off values of SAM were decided based on the 9-point scale, and adapted from Cordon et al. (2013). The set cut-off value for negative outcome was 1-3, for neutral outcome it was 4-6, and for positive outcome 7-9. Table 2 displays the mean valence rating for the original nine scripts.

Table 2

*SAM-mean for the emotional valence response for the nine scripts*

	Morning	Home	Bicycle	Mountain	Run	Gambler	Boy	Date	Birthday
Negative	3.00	2.67	3.00	2.16	3.00	2.50	2.00	2.33	<b>4.00</b>
Positive	8.50	7.16	8.67	7.33	8.00	8.33	7.16	7.00	7.33
Neutral	5.67	<b>6.33</b>	5.83	5.50	5.16	5.16	<b>7.00</b>	5.33	<b>7.00</b>

*Note.* SAM-means that are not within the set cut-off values are in bold face.

The birthday-script, where a young woman celebrates her birthday-party, was omitted, as the mean rating of the neutral outcome of the script was 7.00, which is above the cut-off value for neutral. As the mean for the positive version was only 7.33, the difference between these two was also quite low. The information from the qualitative pilot test supported this decision, as the children generally explained all the photographs as “happy” and positive, as this script takes place in a birthday-situation, which is a situation that can altogether be rated as rather pleasant.

The second script omitted was the boy-script, where a boy visits the grocery store with his mother. This script had a high mean rating for the neutral outcome, which again was 7.00. Here, the mean for the positive outcome was even more similar (7.16). Omitting this script was supported by the qualitative test, as the children stated that the positive and neutral photographs were rather similar, as they both display the boy meeting his friends.

A series of one-way ANOVA were conducted to test if there was a significant difference between the ratings of valence. The children’s scores for valence as indicated on the SAM-assessment were used as the dependent variable. The independent variable was if they rated a negative, positive or neutral outcome. All seven remaining scripts reached significance, Morning script,  $F(2, 17) = 25.37, p = .000, \eta_p^2 = .08$ , Homecoming script,  $F(2, 17) = 23.99, p = .00, \eta_p^2 = .08$ , Bicycle script,  $F(2, 17) = 88.47, p = .00, \eta_p^2 = .09$ , Mountain script,  $F(2, 17) = 34.95, p = .00, \eta_p^2 = .08$ , Run script,  $F(2, 17) = 26.99, p = .00, \eta_p^2 = .07$ , Gambling script,  $F(2, 17) = 32.43, p = .00, \eta_p^2 = .08$ , and Date script,  $F(2, 17) = 24.31, p = .00, \eta_p^2 = .08$ . Post-hoc tests using Bonferroni indicated significant differently ratings of neutral, negative and positive scripts.

The mean ratings for arousal on the different scripts were then further investigated. Means are presented in table 3. The scripts should stir a higher arousal for negative and positive versions of scripts than for neutral versions. The cut off-values for neutral outcomes were set to 1-4, and the cut off-values for positive and negative between 5-9, adapted from Cordon et al. (2013). The homecoming-script did not meet the above-mentioned cut-offs, and



was omitted. Another one-way ANOVA was conducted to explore if any remaining scripts should be omitted based on arousal level. All six remaining scripts were significantly different between neutral versions and both positive and negative versions, Morning script,  $F(2,17) = 3.94, p = .042, \eta_p^2 = .03$ , Bicycle script,  $F(2, 17) = 7.18, p = .006, \eta_p^2 = .05$ , Mountain script,  $F(2, 17) = 10.58, p = .001, \eta_p^2 = .06$ , Run script,  $F(2, 17) = 4.77, p = .025, \eta_p^2 = .04$ , Gambling script,  $F(2, 17) = 10.56, p = .001, \eta_p^2 = .06$ , and Date script,  $F(2,17) = 4.65, p = .002, \eta_p^2 = .04$ . Post-hoc tests using Bonferroni indicated significant different ratings of neutral scripts from both negative and positive scripts, but not a significant difference between ratings for negative and positive scripts.

Table 3

*SAM-mean for the emotional arousal response for the nine scripts*

	Morning	Home	Bicycle	Mountain	Run	Gambler	Boy	Date	Birthday
Negative	6.00	<b>4.17</b>	6.17	7.17	6.50	6.83	5.83	6.50	5.33
Positive	6.83	<b>4.50</b>	7.17	6.50	6.50	6.67	6.50	6.50	<b>4.50</b>
Neutral	3.00	<b>4.17</b>	2.83	3.33	3.00	2.33	<b>5.33</b>	3.00	<b>5.50</b>

*Note.* SAM-means that are not within the set cut-off values are in boldface.

## Results

### Estimating Sample Size

The sample size was estimated based on standard deviation and sample mean from a study using the same paradigm as the present one (Mirandola et al., 2013). Importantly, only means and standard deviation for negative and neutral valence, and not for positive valence and age, was used for the present estimates. This particular choice was based on the fact that one can usually find smaller differences in sample means for valence compared to age (Brainerd et al., 2010). If numbers from data on age had been used, a smaller sample size would have been estimated, and the later analysis might not have reached statistical significance due to the small sample size. As positive valence has not been included in former research with the same paradigm, it was not possible to find any data on this valence. A power of .80 was expected, and the  $p$ -value was set to .05. According to Howells's (2012) appendix, this equals  $\delta = 2.8$ . The estimated needed sample size was thus  $56.3 \approx 57$ .

### Preliminary Analyses

To investigate gender differences in false recognition, an independent-samples t-test was conducted. This was to establish if there was a gender difference in distribution of misses and false alarms. There was not found a significant gender difference for neither false alarms,  $t(130) = 20.75, p = .457, \eta_p^2 = .004$ , nor misses,  $t(130) = 0.91, p = .362, \eta_p^2 = .01$ .

A one-way ANOVA was conducted to explore if the PowerPoint version viewed by the subject did affect overall memory errors in the experiment. The subjects were divided into three groups according to which PowerPoint presentation they watched, and thus which emotional ending to the six different scripts they viewed. All three PowerPoint versions contain two positive endings, two neutral endings and two negative endings, but not for the same scripts. As expected, there was not a statistically significant difference in memory scores depending on a PowerPoint version presented,  $F(2, 128) = 0.59, p = .943, \eta_p^2 = .001$ .

To explore if there was a difference between the two elementary schools regarding false recognition, an independent-samples t-test was conducted with false alarms and misses as dependent test-variables. The participants from school 1 ( $M = 8.05, SD = 5.29$ ) had on average more false alarms than school 2 ( $M = 5.70, SD = 3.89$ ). This difference was

significant,  $t(130) = 2.84, p = .005, \eta_p^2 = .06$ . When looking at differences in misses between the two schools, no statistical difference was revealed.

To see if the differences in false alarms between the two schools could be due to age differences, a new independent t-test was conducted. The average age in school 1 ( $M = 9.38, SD = 1.93$ ) was marginally higher than in school 2 ( $M = 8.76, SD = 1.96$ ). This difference between schools was not significant,  $t(130) = 1.67, p = .098, \eta_p^2 = .02$ . However, with a slight difference between means and a small effect size between the two schools, the differences in total wrong answers were attributed to age differences between schools, and are therefore not included in further analyses.

## Main Analyses

### Proportion Scores

A 2 (picture type; cause vs. gap-filling) X 3 (emotion; positive vs. neutral vs. negative) x 3 (Age; 6-8 vs. 9-10 vs. 11-12 years) repeated measures mixed factorial design was conducted, with picture-type and emotion as within-group factors, and age as a between-group factor. The variables are calculated using proportion scores of misses and false alarms respectively.

No significant main effect of age was found,  $F(2,126) = 2.96, p = .055, \eta_p^2 = .05$ . Post-hoc comparisons using Bonferroni adjustments revealed a trend of children aged 6-8 years ( $M = 37.9, SD = 13.9$ ) having a higher rate of false recognition than children aged 9-10 years ( $M = 33.75, SD = 24.88$ ), with a significance level of  $p = .051$ .

A significant main effect of picture-type was revealed,  $F(1,129) = 88.04, p = .000, \eta_p^2 = .41$ . The rate of false recognition in gap-filling pictures ( $M = 22.20, SD = 15.35$ ) was significantly lower than in cause pictures ( $M = 44.43, SD = 37.71$ ), indicating that children made most mistakes on cause pictures.

A significant main effect of emotion was revealed,  $F(2,128) = 6.51, p = .002, \eta_p^2 = .09$ . Post-hoc comparisons using Bonferroni adjustments revealed that children had significantly lower false recognition of positive scripts ( $M = 20.58, SD = 13.33$ ) compared to negative ( $M = 24.64, SD = 12.88$ ) and neutral ( $M = 25.58, SD = 15.15$ ) scripts. The difference between neutral and negative scripts was not significant.

The above-mentioned main effects were qualified by a significant interaction effect of picture-type and emotion,  $F(2,128) = 7.62, p = .001, \eta_p^2 = .11$ . Post-hoc comparisons using the Bonferroni test adjusted for multiple comparisons revealed that children had significantly

more false alarms for neutral cause pictures ( $M = 50.37$ ,  $SD = 36.8$ ) compared to positive cause pictures ( $M = 35.98$ ,  $SD = 36.18$ ). No other interactions between picture-type and emotion reached significance. Means and standard deviations are presented in table 4.

Table 4

*Mean proportions (and standard deviations) of false recognition in gap-filling and causal errors as a function of emotion*

	Gap-filling	Causal errors
Emotion		
Positive	21.75 (14.87)	35.98 (36.18)
Neutral	21.2 (14.97)	50.37 (36.80)
Negative	23.75 (16.19)	46.59 (40.13)

The interaction between age and picture-type did not reach significance,  $F(2,129) = 0.09$ ,  $p = .914$ ,  $\eta_p^2 = .001$ . The tendency in false alarms for both cause pictures and gap-filling pictures was an age decline in false memories between children aged 6-8 and 9-10 years, but an age increase in false memories between children aged 9-10 and 11-12 years. The means and standard deviations are presented in table 5, and the means are presented in figure 1.

No other significant ( $p > .05$ ) main or interaction effects were revealed.

Table 5

*Mean proportions (and standard deviations) of gap-filling and causal errors (and associated false alarm and miss responses) as a function of age group*

	Gap-filling		Causal errors
	False alarms	Miss	False alarm
Age-group			
6-8 years	14.41 (9.09)	11.98 (6.45)	49.41 (26.14)
9-10 years	13.23 (11.27)	12.29 (9.80)	41.98 (28.70)
11-12 years	14.57 (9.40)	13.13 (7.88)	46.11 (27.98)

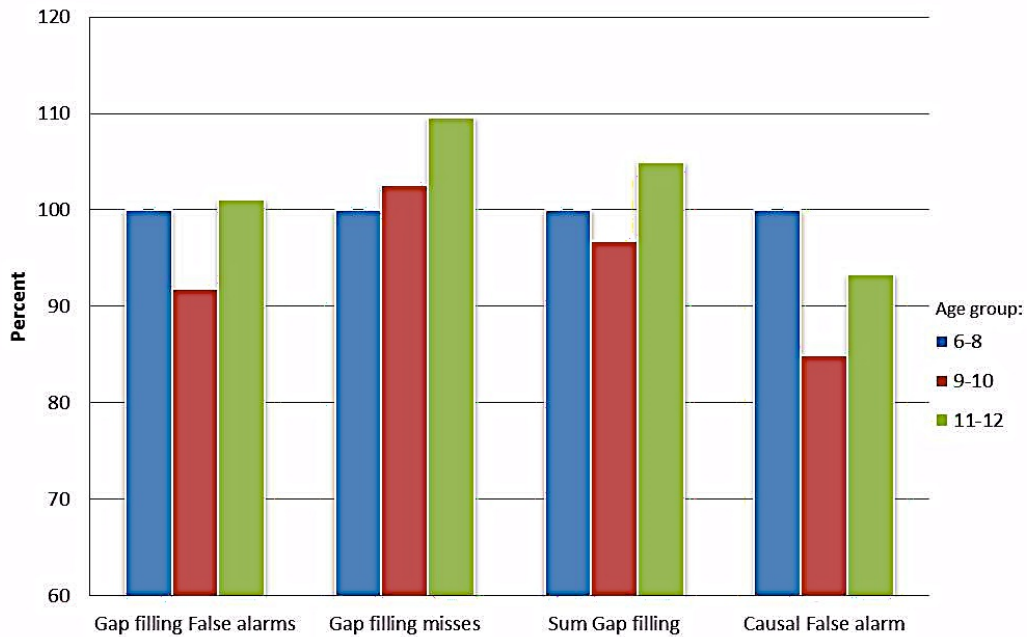


Figure 1. Proportion of memory errors by age group, indexed by age 6-8 years. Age 6-8 is set as the reference category, with a constant score of 100. The other age-groups are compared to the results of age 6-8, indicating, in percent, if they had more or less errors relative to this age-group. The chart begins at 60 percent to better illustrate the trends.

To further explore the interaction between age and false recognition, a one-way ANOVA was conducted to explore developmental trends in type of false answer (false alarm vs. miss) in gap-filling pictures. Cause pictures were not included in the analysis, as only false alarms are possible for this picture-type. There was not found a statistically significant effect of age group for false alarm,  $F(2, 131) = .21, p = .808, \eta_p^2 = .003$  or on miss,  $F(2, 131) = .23, p = .796, \eta_p^2 = .003$ . A trend towards different developmental courses was however indicated. For false alarms, there was no clear age-trend. For misses, there was a linear age-increase, with the youngest children having the lowest proportion of misses. The means and standard deviations are presented in table 5, and the means are presented in figure 1.

### Discrimination Accuracy Index

To investigate memory accuracy as a function of emotion, three  $P_r$ -scores were calculated; two for emotional errors, positive and negative respectively, and one for neutral, using “old” and “new” gap-filling pictures.  $P_r$  scores were calculated by subtracting corrected false alarm- rates from hit-rates.

The  $P_r$  scores were used in a 3 (emotion; positive vs. neutral vs. negative) X 3 (age; 6-8 vs. 9-10 vs. 11-12 years) repeated measures mixed factorial design, with emotion as a within-group factor, and age as a between-group factor.

A main effect on age was found,  $F(2, 129) = 4.64, p = .011, \eta p^2 = .07$ . A Post-hoc comparisons using Bonferroni revealed that the 9-10 year olds ( $M = .28, SD = .11$ ) were significantly better at discriminating between “old” and “new” pictures than the 6-8 year olds ( $M = .22, SD = .14$ ),  $p = .009$ . Further, there was no significant difference between 9-10 year olds and 11-12 year olds ( $M = .24, SD = .13$ ), although the 9-10 year olds had fewer memory errors. There was no significant difference between 6-8 year olds and 11-12 year olds, and no main effect of emotion. No significant interaction effects appeared ( $p > .05$ ), indicating that age does not impact discrimination accuracy regardless of emotions.

Table 6

*Means and standard deviations for  $P_r$ ,  $B_r$ , Hits and False alarms for the recognition-task*

Age-group	6-8 year	9-10 year	11-12 year
$P_r$			
Positive	.23 (.13)	.29 (.11)	.25 (.12)
Neutral	.23 (.12)	.28 (.12)	.22 (.12)
Negative	.20 (.15)	.27 (.11)	.24 (.15)
$B_r$			
Positive	.20 (.12)	.17 (.09)	.17 (.10)
Neutral	.15 (.08)	.12 (.06)	.14 (.06)
Negative	.21 (.13)	.18 (.09)	.18 (.10)
Hits*			
Positive	.39 (.10)	.41 (.09)	.38 (.11)
Neutral	.40 (.09)	.41 (.08)	.38 (.11)
Negative	.38 (.11)	.40 (.08)	.38 (.11)
False alarms*			
Positive	.16 (.11)	.13 (.07)	.13 (.09)
Neutral	.17 (.12)	.14 (.08)	.16 (.07)
Negative	.18 (.13)	.14 (.08)	.15 (.10)

*Note.* Positively, neutrally, and negatively, valenced pictures.  $P_r$  is a discrimination accuracy index, where higher value indicates a better ability to discriminate between «old» and «new» pictures in a recognition memory test.  $B_r$  is a response criterion index set for the participant where a score above 0.5 is considered a liberal response bias, whereas a score below 0.5 is considered conservative. \*Corrected hit- and false alarm-scores are used to calculate  $P_r$  and  $B_r$ .

### **Response Criterion**

A 3 (emotion; positive vs. neutral vs. negative) X 3 (age; 6-8 vs. 9-10 vs. 11-12 years) repeated mixed ANOVA was conducted using the response bias criterion ( $B_r$ ) set for each child when responding to the “old” and “new” gap-filling pictures in the recognition-phase as a dependent factor.

With a significant violation of the assumption of sphericity, the multivariate tests indicated a significant main effect of emotion on gap-filling scores using the  $B_r$  response bias,  $F(2, 128) = 28.81, p = .00, \eta p^2 = .31$ . Bonferroni post-hoc comparisons revealed a higher and more liberal bias for the positive scripts ( $M = .18, SD = .10$ ) compared to the neutral scripts ( $M = .14, SD = .07$ ),  $p = .00$ . There was also a significantly more liberal bias for the negative scripts ( $M = .19, SD = .11$ ) compared to the neutral scripts,  $p = .00$ , indicating that the children were more likely to say “yes” for an emotional script rather than for neutral scripts.

There was no significant difference between the emotional scripts. No other significant main or interaction effects appeared ( $p > .05$ ). This indicates that age does not impact children’s response bias in any of the emotional situations.

## Discussion

When assessing memory using the discrimination accuracy index with corrected hits and false alarms, a significant main-effect of age was found. Children aged 9-10 years were significantly better at discriminating between “old” and “new” pictures compared to children aged 6-8 years. This finding partly confirms the first hypothesis of an age-decrease in memory distortion. However, no significant difference between the 6-8 year olds and the 11-12 year olds emerged. Although not significant, 9-10 year olds were surprisingly slightly better than the 11-12 year olds at discriminating between “old” and “new” pictures.

Children’s gap-filling errors were significantly fewer than causal errors regardless of age. As no age-related increases in memory distortions were revealed, the second hypothesis stating developmental reversals regarding causal errors, and age invariance in gap-filling errors, may be rejected.

Third, we postulated that false recognition would increase if the script was negative, and decrease if the script was positive. Partly supporting the hypothesis, children had a significantly lower rate of memory errors for positive scripts than for both negative and neutral scripts, but not higher for negative than for neutral scripts. A significantly more liberal bias for positive and negative vs. neutral scripts were found. No significant main effect of emotion was found using the discrimination accuracy index, indicating that children’s ability to discriminate between “old” and “new” pictures was not affected by the emotionality of the script.

### **Age Related Performance in Recognition Memory**

The children aged 9-10 years significantly outperformed the younger ones in discriminating between “old” and “new” pictures. This effect was however not significant using proportion scores, although the same trend was evident. This general increase in memory accuracy corroborates previous research on children’s recognition memory (Ackerman, 1984; Lindsay et al., 1991; Davidson & Hoe, 1993). Abstract knowledge structures such as scripts for describing the temporal and causal sequences of events depend partly on language development (Goswami, 2008). In addition to language, children use other symbols such as pictures to encode and communicate their experiences. Symbolic understanding itself develops, and this development is thus a factor in explaining why older children outperform younger children on memory tests. A general notion that seems to be



consistent with the present results is that the development of ability to recognize previously encountered items, and to derive meaning from such items, occurs at the same time as the development of the ability to discriminate between memory traces, even those that are highly similar to previously encountered material (Ghetti et al., 2002; Johnson, Hashtroudi, & Lindsay, 1993).

However, there was no general developmental increase in recognition memory between children aged 9-10 and 11-12 years. Surprisingly, there was a trend indicating older children to make more overall memory errors compared to the 9-10 year olds. This unexpected finding might be due to several reasons.

Recognition has by some been suggested to be a processing skill rather than a cognitive skill (Fagan, 1992), based on the notion that recognition memory seems to be fairly common for animals as well (Vaughan & Greene, 1984). Cognitive development studies might therefore not find much age-development in recognition memory. Brown and Scott (1971) found children between three and five years of age to remember 100 pictures of familiar situations with 98 % accuracy in recognition. Several other experiments have also failed to find a developmental difference in recognition tasks (Caroll, Byrne, & Kishner, 1985; Ghetti et al., 2002; Naito, 1990).

During recognition, the older children seemed more self-conscious, and worried about making memory mistakes compared to the younger children. This worrying may have contributed to them having more false alarms on both gap-filling pictures and causal pictures compared to the 8-9 year olds. The response bias did not show any significant age-differences, thus the lack of an age difference between the younger and oldest children in the present study cannot be explained by the children's response criterion. Another reason for this unexpected trend might be the somewhat lower sample of children aged 11-12 years (56: 6-8 year olds vs. 43: 9-10 year olds vs. 33: 11-12 year olds).

### **False Memories and Age**

The present results showed no clear signs of the anticipated developmental reversals in false memories for causal pictures in particular. This finding is thus failing to support the FTT and previously reported research showing developmental reversals in false memory formation (Brainerd et al., 2004; see Brainerd & Reyna, 2012, for a review; Brainerd, Reyna, Ceci, & Holliday, 2008; Holliday et al., 2011; Howe, 2007; Lyons et al., 2010; Metzger et al., 2008). FTT predicts that false memory effects may increase with age as a consequence of

developmental differences in the formation of gist traces (Ghetti et al., 2002). Because children get better at extracting the gist from what they see with age, the chance of creating false memories as a result of studying pictures or word-lists converging on a theme (i.e. gist) should also increase, causing the younger children to show reduced false memory for related distractors. However, gap-filling pictures should be exempt of this effect given that they are more reliant on verbatim memory compared to gist. Thus, the present findings support this notion by revealing non-significant age differences in false alarms and misses on gap-filling pictures.

Using the same Backward Causal-Inference Paradigm with children, Lyons and colleagues (2010) found developmental reversals in recognition memory, indicated by nine year olds being significantly more likely to make causal errors than six year olds. In the present study there was no significant difference in age on causal errors, but children aged 6-8 eight years made more errors in general than children aged 9-10 years. Although the results of the present study are not significant, similar trends as reported by Lyons and colleagues (2010) can be detected, in that the oldest children aged 11-12 years conducted more causal errors than the children aged 9-10 years.

There may be several reasons for why the results in the present study differ from that of Lyons and co-workers (2010), and other developmental studies done with the aim of detecting developmental reversals in false memories (for a review, see Brainerd & Reyna 2012). In the present study, children viewed six different scripts, whereas in Lyons and colleagues (2010) study, the children only viewed four scripts. Watching six different scripts might be too much for the youngest children, causing them to make more errors, and therefore masking the predicted developmental reversals.

Another plausible reason might be the manipulation of valence, which was varied within subjects. Three different outcomes of emotionality might be too much to grasp for the youngest children. The difference in results might be due to the developing ability to focus attention and grasp variety in stimuli instead of actually measuring developmental differences in false memory for emotional events. Older children become more aware and prepared for variation, whereas younger children have the tendency to overlook variation, and need more experience to comprehend and take variation into account (von Tetzchner, 2012). The age range included in the present study might have been too young and all children, regardless of age, might have been confused by the different valences and the number of story-scripts which in turn resulted in the lack of the expected developmental effect of picture type.

Another reason for the lack of developmental reversals could be that the causal pictures for each script are not clearly enough related to the previously seen effect, and might therefore rather resemble a gap-filling picture, which is not predicted to cause any age-difference. As before mentioned, former research has hypothesized the age-increase in memory errors to be due to older children's and adult's ability to connect the meaning of studied words and items, leading the critical lure to get coded along with the items studied during the encoding phase (Roediger & McDermott, 1995). This effect might not be present if the participants do not perceive the causal pictures as a cause to the event they just saw, and these pictures might therefore not trigger any causal explanation to the event during recognition. However, the present study did find significant differences in errors between the causal pictures and the gap-filling pictures. The children made significantly more causal errors as opposed to gap-filling errors, indicating that there was an actual qualitative difference between the picture-types.

The non-reversal effect revealed, might also be explained in line with previous critique of the DRM-lists method; that age-increases in memory errors potentially are due to age-related increases in children's ability to detect relations among items on a word-list and not an direct effect of memory development (e.g. Brainerd, Reyna, & Ceci, 2008; Howe, 2008; Lyons et al., 2010). Therefore, an age-increase in memory errors might not be as evident when using other methods like the Backward Causal-Inference Paradigm that, in contrast to the DRM-paradigm, does not rely on the ability to detect the overall theme of a word-list across multiple items, making it easier for younger children to remember previously encountered stimuli.

Young children, compared to older children and adults, seem to be more dependent on cues to remember (Qin, Quas, Redlich, & Goodman, 1997; Reese, 2008). The pictorial stimuli given in a recognition phase may work as cues for the younger children, allowing them to score higher in recognition tests. Therefore, one might expect to find greater developmental differences in memory distortion using a recall test where no stimuli (e.g. pictures) are present to work as cues as opposed to a recognition test. This is in line with previous research indicating that the discrepancy between different age groups is larger when using recall tests compared to recognition tests (Ceci & Bruck, 1993; Ghetti et al., 2002).

## **Divergent Influences of Emotionality on Memory**

Partly in line with our hypothesis and previous research (Brainerd et al., 2010; Brainerd, Stein et al., 2008), positive scripts in the present study yielded significantly fewer false alarms than negative and neutral scripts. However, negative and neutral scripts did not significantly differ.

When using the discrimination accuracy index, no effect of emotion on ability to discriminate between “old” and “new” photographs was revealed. In line with the present results, several other studies that found an effect of emotion on false memory, did not report a significant effect of emotion on memory when using discrimination accuracy indices (e.g., Johanson, Mecklinger, & Treese, 2004; Kensinger & Corkin, 2003, fifth experiment; Maratos, Allen, & Rug, 2000). Due to these inconsistent findings based on the way in which memory is measured, the impact of emotion on memory has been suggested to be due to other factors, such as response bias (Dougal & Rotello, 2007).

In the present study, a significantly more liberal response bias for both positive and negative photographs compared to neutral photographs was found, indicating that the children were more likely to say “yes” when responding to emotional as opposed to neutral photographs. No significant difference was found between response bias for negative and positive photographs. This indicates that even though the children were most likely to conduct false alarms on neutral distractors, they were more likely to accept either positive or negative distractors as previously seen when in an uncertain state. Previous research has indicated that as emotion changes from neutral to negative, a more liberal bias for accepting distractors as seen emerges in both children (Augusti & Melinder, 2012; Pollak et al., 2000) and adults (Johanson et al., 2004; Kapucu, Rotello, Ready, & Seidl, 2008). Dougal and Rotello (2007) therefore postulated the heightened level of both true and false memory for emotional stimuli to be a factor of response bias (i.e., a more liberal bias for emotional stimuli). This heightened liberal bias for emotional content rather than neutral content has been proposed to be a factor of an evolved ability to make faster decisions in more emotional situations (Phelps & Sharot, 2008).

Several variables could explain why the liberal response bias, discrimination accuracy index, and proportion false alarms are differently affected by emotion. The small variance in arousal between different valences could for instance explain the somewhat atypical results for negative scripts. Valence is significantly different as a function of emotion (neutral, negative, and positive), whereas arousal is only differing between neutral and emotional

(negative and positive) scripts. Thus, as only valence is consistently different as a function of emotion of the scripts, it is not possible to distinguish if the effects of emotion are due to arousal or valence. The pilot-test is quite comprehensive in order to address this shortcoming. All scripts included in the experiment were significantly different between the three valences, which was qualified by moderate to large effect sizes for all scripts. Although the scripts were significantly different between levels of arousal for neutral vs. positive or negative stimuli, the effect sizes were small to moderate. This suggests that the scripts might not be arousing enough, as the mean arousal ratings revealed in the pilot test were towards the lower accepted values for negative and positive outcomes, and towards the higher cut-offs for neutral outcomes. The differences in arousal between valences are thus smaller than the differences in valence, and the arousal ratings for negative and positive pictures are not significantly different and thus not consistently manipulated. As valence and arousal are postulated to influence memory through different mechanisms and in different strength (Kensinger, 2004), this could explain the non-significant differences between negative and neutral stimuli. Previous research has found effects of negative valence only leading to poorer memory when high arousing (Kensinger, 2004). Higher arousal is thought to have a more automatic effect on memory, whereas the effect of positive and negative valence on memory might be due to conscious encoding strategies (Kensinger, 2004). As encoding strategies still develop in the age-range included in the present study (Goswami, 2008), the effects of differently valenced material might be lower in children of this age.

However, in a study conducted with children, Brainerd et al. (2010) found negative valence to elevate false alarms regardless of level of arousal. This indicates that negative valence alone should be enough to increase false alarms, an effect that should increase further with age (Rivers et al., 2008). As arousal intensifies the effect of valence (Brainerd et al., 2010), a smaller effect should be found for lower arousing negative stimuli. The smaller differences in arousal as well as the young age of participants might therefore explain the unexpected results for negative stimuli in the present study.

In addition, the effect of emotion was qualified by an emotion by picture-type interaction. This interaction does to a degree confirm the validity of the arousal and valences in the scripts, as the present study replicates the typical heightened effect of emotionality on causal errors, as neutral stimuli compared to positive stimuli were significantly associated with more causal errors. However, this heightened effect was only found for positive scripts, and not for negative scripts, which did not differ from neutral scripts. Previous research conducted using the same paradigm as the present one has indicated a significant effect of

emotion on cause pictures, but not on gap-filling pictures, for both children (Mirandola et al., 2014) and young adults (Toffalini, Mirandola, Drabika, Melinder, & Cornoldi, 2014). The effect of emotionality was suggested to be higher on cause pictures as these are more tightly bound to the emotionality compared to gap-filling pictures, and gap-filling pictures thus being more indirectly influenced by emotionality (Brainerd, Stein et al., 2008; Mirandola et al., 2014). In the only child study using the same paradigm and including emotions as a within-group variable, typically developing children were found to have a higher proportion of false alarms for negative causes than for neutral causes (Mirandola et al., 2014). As this was found in a rather small sample ( $N=19$ ), the difference in results compared to the present study is particularly interesting as there should be enough power in the present study to detect the same interaction if it in fact existed. These divergent results do nevertheless point towards an effect of emotion on cause pictures, an effect that was not significant for gap-filling pictures.

In all, the overall trend of non-difference between neutral and negative stimuli is conflicting with the majority of other research, using the same (Mirandola et al., 2014) and other (for a review, see Brainerd and Reyna, 2012) paradigms. It does however fit with the typical developmental trends for negative stimuli as the high level in false alarms for negative stimuli has been found to develop with age (Brainerd et al., 2010), with the higher degree of false memories for negative stimuli relative to positive stimuli emerging between the ages of seven and 11 years. The robust larger effect of negative stimuli on false alarms (Brainerd, Stein et al., 2008) might therefore not be developed yet in the present sample of children, as the children's mean age is quite young. Although there was no significant interaction between age and emotionality on false alarm, there were different developmental trends for the emotions. The difference in proportion of false alarms between the ages of 6-8 years and 9-10 years was larger for negative than for positive stimuli, and the oldest age group had the most mistakes on negative stimuli, fitting with the theory of an age-increase in false alarms for negative stimuli (Brainerd et al., 2010). The developmental increase in false alarms for negative stimuli could thus be postulated to not yet have emerged in the children tested in the present study. The tendencies of a more liberal bias for emotional stimuli, combined with the higher degree of false errors in negative and neutral scripts, show a particularly important tendency for false memories in negative situations. This supports the previous found tendencies of negative stimuli not protecting against false memory, but rather the opposite (see Brainerd & Reyna, 2012, for a review).

It is important to keep in mind that this is at the present date the first study to research both negative and positive vs. neutral effects on gap-filling and causal interference with a developmental perspective. The untypical finding for negative stimuli and developmental reversals might therefore be a factor of criterion validity –previous research might measure slightly different mechanisms than the present one.

### **Limitations of the Present Study**

The present experiment is a laboratory study, and is therefore somewhat limited when applied to real-life situations. From an applied perspective, the results from this study are more applicable to memory for passive bystanders. However, the scripts display everyday-situations, with events such as falling in a race or finding a spider. These events are less arousing than witnessing a robbery or other criminal situations. Particularly when concerning children, ethical considerations are of great importance. This limits both number of morally accepted ways of conducting experiments, and what stimuli can be included. On the other hand, a laboratory study allows for controlled investigation of theoretically driven phenomena by minimizing the effect of potentially confounding variables.

The scripts displayed, such as dating, are not as common or interesting for all age groups. What people remember of a picture is influenced by how they encode it (Schacter, 1996), and people typically remember different aspects of stimuli based on their own background and interests. Children participating in the present study highlighted different aspects of the scripts displayed. When talking after the experiment, several children explained how they linked aspects of the pictures to their own lives. Some also asked to see the pictures again, whereas some children were clearly less interested. It is important to note that all the scripts in the present study display only adults, possibly causing the scripts to be less relevant and interesting for the children.

Another possible problem with the study could be the number of scripts, as six scripts with a total of three different emotions might demand too much of the children's attention. Particularly for the younger children, the effects of the emotions might cause a spill over-effect that in turn might have affected the results.

## **Future Directions**

Larger effects of developmental reversals have been found in studies also including older participants (Mirandola et al., 2013), indicating the importance of investigating the effect of emotion on false memory with an even larger age-span. This could be particularly interesting regarding the effects of positive stimuli, as little is known about the developmental trajectories of the effect this induces on false memory. A larger age-span could also be interesting regarding response bias, as a developmental effect on liberal bias between childhood and adolescence has been indicated (Cordon et al., 2013), with only adults eliciting a significantly more liberal response bias for negative than for neutral pictures. The present study did not find an interaction between age and emotionality for response criterion; however, this might be due to the limited age range included in the present study. Thus, more developmental research on the interaction between response criterion and all three emotions would be interesting, as there are no overall clear tendencies on the effect of age on response criterion, which has been postulated to be behind the effect of emotion on false memory (Dougal & Rotello, 2007).

From an applied perspective, the tendencies regarding developmental reversals are important to investigate further, as there is still not consensus on exactly how much and in which ways development affects false memory (Magnussen, 2004). The importance of more research on the subject is stressed by the divergent opinions and understanding of how false memories operate, both by the general population (Magnussen et al., 2006), and by experts such as judges (Magnussen, Melinder, Stridbeck, & Raja, 2010; Magnussen et al., 2008), and psychologist (Magnussen & Melinder, 2012) as well as by researchers (Patihis, Ho, Tingen, Lilienfeld, & Loftus, 2013; see Sabbagh, 2009, for a review). The opinion and knowledge these groups have for false memory could be particularly important in a courtroom, where they act as deputies, judges, or expert witnesses. To be applicable in a court of law, research in false memory must simulate relevant situations. This paradigm should aim to include material that is as close to children's reality as possible, thus increasing the generalizability of future studies. Future studies should therefore aim to improve the child-appropriateness of the scripts, with scripts displaying situations that are more typical and interesting for all age groups of children, and primarily using children as actors in the scripts.



## **Concluding Remarks**

The aim of the present study was to investigate how emotion influences memory errors, from a developmental-perspective. The impact of age, picture-type, and emotionality on recognition memory was studied with the Backwards Causal-Inference Paradigm. Results indicated that memory performance increases from the ages of 6-8 to 9-10 years, whereas no further age increases or decreases were detected. In addition, a lower degree of memory errors was detected for positive stimuli compared to neutral and negative stimuli. A higher liberal bias was indicated for both positive and negative scripts compared to neutral scripts.

The dispute regarding formation of memory distortions seems to be endless, particularly concerning children. More controlled research is desirable to untangle this dispute, from both an applied and a scientific perspective. The present study attempts to clarify this debate, specifically regarding the impact of both negatively and positively perceived events, and the similarities and differences between the two. As positive events are indicated to have a protective effect against memory errors, with the same tendency not present for negative events, this suggests a qualitative difference in how emotions affect memory in children.

## References

- Ackerman, B. P. (1984). Item-specific and relational encoding effects in children's recall and recognition memory for words. *Journal of Experimental Child Psychology*, 37(3), 426-450.
- Anastasi, J. S., & Rhodes, M. G. (2008). Examining differences in the levels of false memories in children and adults using child-normed lists. *Developmental Psychology*, 44(3), 889- 894. doi: 10.1037/0012-1649.44.3.889
- Augusti, E. M., & Melinder, A. (2013). The effect of neutral and negative colour photographs on children's item directed forgetting. *European Journal of Developmental Psychology*, 10(3), 378-391. doi: 10.1080/17405629.2012.686741
- Backs, R. W., da Silva, S. P., & Han, K. (2005). A comparison of younger and older adults' self-assessment manikin ratings of affective pictures. *Experimental aging research*, 31(4), 421-440.
- Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(1), 197- 215. doi: 10.1037/0278-7393.22.1.197
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.
- Brainerd, C. J., Forrest, T., Karibian, D., & Reyna, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology*, 42(5), 962-979.
- Brainerd, C. J., Holliday, R., & Reyna, V. F. (2004). Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child development*, 75(2), 505-522.
- Brainerd, C. J., Holliday, R., Reyna, V. F., Yang, Y., & Toglia, M. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology*, 107(2), 137-154.
- Brainerd, C.J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review*, 24(4), 396-439.
- Brainerd, C. J., & Reyna, V. F. (2012). Reliability of children's testimony in the era of developmental reversals. *Developmental Review*, 32(3), 224-267.

- Brainerd, C. J., Reyna, V. F., & Ceci, S. (2008). Developmental reversals in false memory: a review of data and theory. *Psychological bulletin*, *134*(3), 343- 382. doi: 10.1037/0033-2909.134.3.343
- Brainerd, C. J., Reyna, V. F., Ceci, S., & Holliday, R. (2008). Understanding developmental reversals in false memory: Reply to Ghetti (2008) and Howe (2008). *Journal of Experimental Child Psychology*, *107*(2), 137-154
- Brainerd, C. J., Reyna, V. F., & Forrest, T. (2002). Are Young Children Susceptible to the False–Memory Illusion? *Child development*, *73*(5), 1363-1377.
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. *Advances in child development and behavior*, *28*, 41-100.
- Brainerd, C. J., Stein, L., Silveira, R., Rohenkohl, G., & Reyna, V. F. (2008). How does negative emotion cause false memories? *Psychological Science*, *19*(9), 919-925.
- Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*(2), 81-129.
- Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. New York: Oxford University Press.
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*(1), 73-99.
- Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual review of psychology*, *50*(1), 419-439.
- Buchanan, T. W., & Adolphs, R. (2002). The role of the human amygdala in emotional modulation of long-term declarative memory. *Advances in Consciousness Research*, *44*, 9-34.
- Burgwyn-Bailes, E., Baker-Ward, L., Gordon, B. N., & Ornstein, P. A. (2001). Children's memory for emergency medical treatment after one year: The impact of individual difference variables on recall and suggestibility. *Applied Cognitive Psychology*, *15*(7), 25-48.
- Burrell, L. V. (2013). *Children's earwitness memory: The influence of age and emotionality*. Oslo: University of Oslo.
- Carroll, M., Byrne, B., & Kirsner, K. (1985). Autobiographical memory and perceptual learning: A developmental study using picture recognition, naming latency, and perceptual identification. *Memory & cognition*, *13*(3), 273-279.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: a historical review and synthesis. *Psychological bulletin*, *113*(3), 403-439.

- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington: American Psychological Association.
- Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis*, *42*(4), 304-320.
- Christianson, S.-Å. (1992). Emotional stress and eyewitness memory: a critical review. *Psychological bulletin*, *112*(2), 284-309. doi: 10.1037/0033-2909.112.2.284
- Cooper, L. A., & Schacter, D. L. (1992). Dissociations between structural and episodic representations of visual objects. *Current Directions in Psychological Science*, *1*(5), 141-146.
- Cordon, I. M., Melinder, A., Goodman, G. S., & Edelstein, R. S. (2013). Children's and adults' memory for emotional pictures: Examining age-related patterns using the Developmental Affective Photo System. *Journal of Experimental Child Psychology*, *114*(2), 339-356.
- Cordon, I. M., Pipe, M. E., Sayfan, L., Melinder, A., & Goodman, G. S. (2004). Memory for traumatic experiences in early childhood. *Developmental Review*, *24*(1), 101-132. doi: 10.1016/j.dr.2003.09.003
- Davidson, D., & Hoe, S. (1993). Children's Recall and Recognition Memory for Typical and Atypical Actions in Script-Based Stories. *Journal of Experimental Child Psychology*, *55*(1), 104-126.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of experimental psychology*, *58*(1), 17-22.
- Dewhurst, S. A., Pursglove, R. C., & Lewis, C. (2007). Story contexts increase susceptibility to the DRM illusion in 5- year- olds. *Developmental Science*, *10*(3), 374-378.
- Dougal, S., & Rotello, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review*, *14*(3), 423-429.
- Fagan, J. F. (1992). Intelligence: A theoretical viewpoint. *Current Directions in Psychological Science*, *1*(3), 82-86.
- Fivush, R., Hazzard, A., McDermott Sales, J., Sarfati, D., & Brown, T. (2003). Creating coherence out of chaos? Children's narratives of emotionally positive and negative events. *Applied Cognitive Psychology*, *17*(1), 1-19.
- Gallo, D. (2013). *Associative illusions of memory: False memory research in DRM and related tasks*. New York: Psychology Press.

- Ghetti, S., & Angelini, L. (2008). The Development of Recollection and Familiarity in Childhood and Adolescence: Evidence From the Dual- Process Signal Detection Model. *Child development, 79*(2), 339-358.
- Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: age, distinctiveness, and subjective experience. *Developmental Psychology, 38*(5), 705-718.
- Goswami, U. (2008). *Cognitive development: the learning brain*. Hove: Psychology Press.
- Grimm, M., & Kroschel, K. (2005). *Evaluation of natural emotions using self assessment manikins*. Paper presented at the Automatic Speech Recognition and Understanding, 2005 IEEE Workshop. doi: 10.1109/ASRU.2005.1566530
- Hannigan, S. L., & Tippens Reinitz, M. (2001). A demonstration and comparison of two types of inference-based memory errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(4), 931.
- Holliday, R. E., Brainerd, C. J., & Reyna, V. F. (2011). Developmental reversals in false memory: Now you see them, now you don't! *Developmental Psychology, 47*(2), 442-449. doi: 10.1037/a0021058
- Howe, M. L. (2007). Children's emotional false memories. *Psychological Science, 18*(10), 856-860.
- Howe, M. L. (2008). What is false memory development the development of? Comment on Brainerd, Reyna, and Ceci (2008). *Psychological Bulletin, 134*(5), 768-772. doi: 10.1037/0033-2909.134.5.768.
- Howe, M. L., Candel, I., Otgaar, H., Malone, C., & Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory, 18*(1), 58-75.
- Howe, M. L., Cicchetti, D., Toth, S. L., & Cerrito, B. M. (2004). True and false memories in maltreated children. *Child development, 75*(5), 1402-1417.
- Howell, D. (2012). *Statistical methods for psychology*. Belmont: Wadsworth.
- Johansson, M., Mecklinger, A., & Treese, A.-C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience, 16*(10), 1840-1853.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin, 114*(1), 3-28.

- Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1409-1419. doi: 10.1037/0278-7393.20.6.1409
- Kapucu, A., Rotello, C. M., Ready, R. E., & Seidl, K. N. (2008). Response bias in "remembering" emotional stimuli: a new perspective on age differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(3), 703-711.
- Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, 15(4), 241-252.
- Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & cognition*, 31(8), 1169-1180.
- Kensinger, E. A., & Schachter, D. L. (2008). Memory and Emotion. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 601-617). New York: Guilford.
- Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, 52(3), 297-318.
- Loftus, E. F. (1979). The malleability of human memory. *American Scientist*. 67(3), 312-320.
- Lyons, K. E., Ghetti, S., & Cornoldi, C. (2010). Age differences in the contribution of recollection and familiarity to false- memory formation: a new paradigm to examine developmental reversals. *Developmental Science*, 13(2), 355-362.
- Magnussen, S. (2004). *Vitnepsykologi: pålitelighet og troverdighet i dagligliv og rettssal*. Oslo: Abstrakt forlag.
- Magnussen, S., Andersson, J., Cornoldi, C., De Beni, R., Endestad, T., Goodman, G. S., . . . Zimmer, H. (2006). What people believe about memory. *Memory*, 14(5), 595-613. doi: 10.1080/09658210600646716
- Magnussen, S., & Melinder, A. (2012). What psychologists know and believe about memory: A survey of practitioners. *Applied Cognitive Psychology*, 26(1), 54-60.
- Magnussen, S., Melinder, A., Stridbeck, U., & Raja, A. Q. (2010). Beliefs about factors affecting the reliability of eyewitness testimony: A comparison of judges, jurors and the general public. *Applied Cognitive Psychology*, 24(1), 122-133.
- Magnussen, S., Wise, R. A., Raja, A. Q., Safer, M. A., Pawlenko, N., & Stridbeck, U. (2008). What judges know about eyewitness testimony: A comparison of Norwegian and US judges. *Psychology, Crime & Law*, 14(3), 177-188.

- Maratos, E. J., Allan, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, 38(11), 1452-1465.
- Metzger, R. L., Warren, A. R., Shelton, J. T., Price, J., Reed, A. W., & Williams, D. (2008). Do children "DRM" like adults? False memory production in children. *Developmental Psychology*, 44(1), 169-181
- Mirandola, C., Losito, N., Ghetti, S., & Cornoldi, C. (2014). Emotional false memories in children with learning disabilities. *Research in developmental disabilities*, 35(2), 261-268.
- Mirandola, C., Paparella, G., Re, A., Ghetti, S., & Cornoldi, C. (2012). Children with ADHD symptoms are less susceptible to gap-filling errors than typically developing children. *Learning and Individual Differences*, 22(6), 896-900.
- Mirandola, C., Toffalini, E., Grassano, M., Cornoldi, C., & Melinder, A. (2013). Inferential false memories of events: Negative consequences protect from distortions when the events are free from further elaboration. *Memory*, 22(5), 451-461.
- Naito, M. (1990). Repetition priming in children and adults: Age-related dissociation between implicit and explicit memory. *Journal of Experimental Child Psychology*, 50(3), 462-484.
- Nelson, C. A., De Haan, M., & Thomas, K. M. (2006). *Neuroscience of cognitive development: the role of experience and the developing brain*. Hoboken, N.J.: John Wiley.
- Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General*, 129(2), 242-261
- Otgaar, H., Peters, M., & Howe, M. L. (2012). Dividing attention lowers children's but increases adults' false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(1), 204-210.
- Patihis, L., Ho, L. Y., Ting, I. W., Lilienfeld, S. O., & Loftus, E. F. (2013). Are the "Memory Wars" Over? A Scientist-Practitioner Gap in Beliefs About Repressed Memory. *Psychological Science*, 25(2), 519-530. doi: 10.1177/0956797613510718
- Perner, J., & Ruffman, T. (1995). Episodic memory and autoevident consciousness: developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology*, 59(3), 516-548.
- Phelps, E. A., & Sharot, T. (2008). How (and why) emotion enhances the subjective sense of recollection. *Current Directions in Psychological Science*, 17(2), 147-152.

- Pollak, S. D., Cicchetti, D., Hornung, K., & Reed, A. (2000). Recognizing emotion in faces: developmental effects of child abuse and neglect. *Developmental Psychology, 36*(5), 679-688.
- Qin, J., Quas, J. A., Redlich, A. D., & Goodman, G. S. (1997). Children's eyewitness testimony: Memory development in the legal context. In N. Cowan & C. Hulme (Eds.), *The development of memory in childhood* (pp. 301-341). New York: Psychology Press.
- Reese, E. (2008). The development of autobiographical memory: origins and consequences. *Advances in child development and behavior, 37*, 145-200.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1-75.
- Reyna, V. F., & Titcomb, A. L. (1997). Constraints of the suggestibility of Eyewitness Testimony: A Fuzzy-Trace Theory Analysis. In D.G. Payne & F.G. Conrad (Eds.), *Intersections in basic and applied memory research* (pp. 157-171). New Jersey: Lawrence Erlbaum Associates Publishers.
- Rivers, S. E., Reyna, V. F., & Mills, B. (2008). Risk taking under the influence: A fuzzy-trace theory of emotion in adolescence. *Developmental Review, 28*(1), 107-144.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803- 814. doi: 10.1037/0278-7393.21.4.803
- Sabbagh, K. (2009). *Remembering our childhood: How memory betrays us*. New York: Oxford University Press.
- Schacter, D. L. (1996). *Searching for memory: the brain, the mind, and the past*. New York: BasicBooks.
- Schooler, J. W., & Eich, E. (2000). Memory for emotional events. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 379-392). New York: Oxford University Press.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34-50.
- Toffalini, E., Mirandola, C., Drabik, M. J., Melinder, A., & Cornoldi, C. (2014). Emotional negative events do not protect against false memories in young adults with depressive-anxious personality traits. *Personality and Individual Differences, 66*, 14-18.

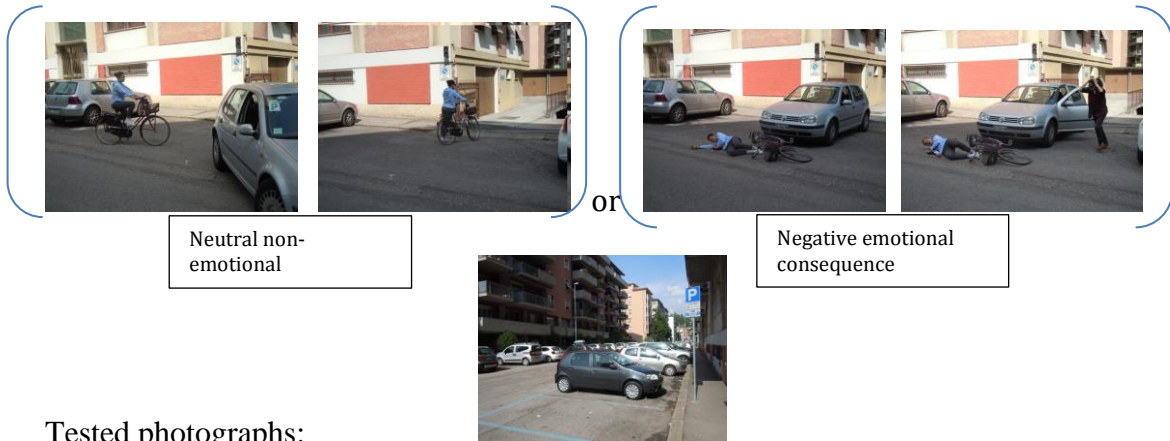


- Vaughan, W., & Greene, S. L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes*, 10(2), 256-271.
- von Tetzchner, S. (2012). *Utviklingspsykologi*. Oslo: Gyldendal Akademisk.
- Wright, R. (1994). *The moral animal: evolutionary psychology and everyday life*. New York: Vintage Books.
- Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and cognition*, 5(4), 418-441.

# Appendix

## Example of the Bicycle Script

### Encoded photographs:



### Tested photographs:



### TARGETS



### DISTRACTORS

### CAUSE