Acknowledgement: This is the accepted author's version of this paper. The published paper is available at http://www.tandfonline.com/doi/abs/10.1080/15248372.2013.848871

Title:

Children's and adults' ability to build online emotional inferences during comprehension of audiovisual and auditory texts

Authors:

Anna Katharina Diergarten (University of Würzburg)

Gerhild Nieding (University of Würzburg)

Abstract

Two studies examined inferences drawn about the protagonist's emotional state in movies (study 1) or audiobooks (study 2). Children aged 5, 8 and 10, and adults took part. Participants saw or heard 20 movie scenes or sections of audio books taken or adapted from the TV show *Lassie*. An online measure of emotional inference was designed that assessed the ability of the participants to understand the main protagonist's emotional state. The participants' emotional knowledge and media literacy were assessed as further variables. The results of the studies provide evidence that children from the age of 5 build emotional inferences when both watching movies and listening to audiobooks. A developmental trend exists with regard to the precision of the emotional inferences. Media literacy and emotional knowledge differed in terms of their influence on the ability to generate inferences, which was dependent on the age of the participant and the presentation mode.

Keywords

Children, development, emotional inferences, emotional knowledge, media literacy, text comprehension

Introduction

Narrative media and emotions are inseparably linked to each other. Emotions are the basic essentials that make books page-turners or movies must-sees, but how can we come to understand the emotions experienced by the various characters? This question belongs to a corpus of text comprehension research that examines how we understand not only emotional information but any sort of information provided by a text. It concerns the nature of the mental representations that media users construct during comprehension. These theories concerning inference generation during narrative text comprehension (Graesser, Singer, &

Trabasso, 1994; Graesser & Zwaan, 1995) assume that the elaboration of readers' representations varies from representations of the text's surface structure to very sophisticated representations containing information that is not even explicitly mentioned in the text (Fletcher, 1994; Graesser, Millis, & Zwaan, 1997; Kintsch, 1998). The latter representations are called "situation models" (van Dijk & Kintsch, 1983). By constructing various kinds of inferences, readers combine information given in the text with their own world knowledge in order to form a sophisticated representation of the state of affairs described or implied by the text. Inferences are therefore presumably embedded in the situation model (Schmalhofer, McDaniel, & Keefe, 2002). According to the event-indexing model (Zwaan, Langston, & Graesser, 1995), readers organize their mental representation along five dimensions: space (the spatial region in which the situation is located), time (the time frame of the given situation), protagonist (which people are involved in this situation), causality (the causal connection of the current situation with prior events), and intentionality (the protagonist's goals). In this article we are concerned with inferences about the protagonist's emotional state.

Emotional Inferences

Recent discussions have deliberated about whether the protagonist's emotion has a dimension of its own (Therriault & Rinck, 2007) or whether it belongs to one of the other dimensions. Gernsbacher (1995) suggests that it is part of the intentionality dimension, since achievement or non-achievement of goals often leads to emotions. On the other hand, Rapp, Gerrig, and Prentice (2001) assume that it belongs to the protagonist's dimension, because the kind of emotion elicited in a specific situation depends on personal dispositions and traits.

The work by Gernsbacher and her colleagues (Gernsbacher, Goldsmith, & Robertson, 1992; Gernsbacher & Robertson, 1992; Gernsbacher, Robertson, Palladino, & Werner, 2004) showed that adult readers construct emotional inferences while reading fictional texts. There

is also evidence for the claim that these inferences are generated online, i.e. during the comprehension process (Gernsbacher, Hallada, & Robertson, 1998; Graesser, Singer, & Trabasso, 1994; Molinari et al., 2009). The commonly used measures for emotional inferences are reading latencies, as in Gernsbacher et al.'s study (1992) in which participants read a story which described certain circumstances that regularly invoked a specific emotion in the main character. The story was presented on a computer screen. After the participants had read the story a single sentence appeared on the screen describing the character's emotion (for example, "it would be weeks before Tom's guilt subsided"; Gernsbacher et al., 1992, p. 93). This sentence either matched or mismatched the character's emotion in the story. Participants were asked to press a key as soon as they had finished reading the sentence. Shorter reading times for matching sentences compared with mismatching ones were interpreted as an indicator for emotional inferences, since the sentence matched the reader's situation model, which facilitated reading.

Another research question concerns the specificity of emotional inferences, that is, whether the elicited inference concerns only the emotional valence or a more differentiated assessment. If, for example, we read, see or hear that someone is smiling, that does not provide a lot of information beyond the fact that the person is probably experiencing a pleasant feeling. On the other hand, if we are given the information that the person has just been given a present, we can conclude that the person's emotion is very probably happiness; if we hear that it is the reward for a very good achievement, we would conclude that the person is probably proud as well. Ohler, Nieding, and Töpper (2002) have shown that adults' emotional inferences go beyond the valence of the protagonist's emotion. For instance, they differentiate between the emotions joy and pride, as in the given example above. On the other hand, there are no differences for very similar emotions such as "depressed" and "miserable" (Gygax, Oakhill, & Garnham, 2003). How specifically the emotion is represented in the situation model seems to depend on several factors, such as the sufficiency of background

information (Molinari et al., 2009) and the reader's emotional knowledge (Gernsbacher & Robertson, 1992).

So far this research methodology has not been extended to children and therefore the developmental course of constructing emotional inferences during text comprehension is unclear. There are, however, a few studies on the development of situation models that show that children, starting around 4 or 5 years of age, build profound situation models, meaning that they construct various kinds of inferences that add up to a complex mental representation of the situation. These inferences concern not only perceptible information (e.g. the protagonist's movement (Fecica & O'Neill, 2010; Ziegler, Mitchell, & Currie, 2005) or location (Nieding, 2006)), but also the protagonist's mental state. The latter type of inferences are also built by children as young as 5 years and concern, for instance, the protagonist's goals (O'Neill & Shultis, 2007; Unsöld, 2008) and other motivational factors influencing his or her actions (Fecica & O'Neill, 2010). Although there has so far not been any research on children's online emotional inferences, many studies have examined emotional understanding in children, as will be summarized in the following paragraph.

Children's Emotional Knowledge

Very young children can only divide emotions into two broad categories – feeling good or feeling bad. The older they get, the more precise their ability to make a distinction becomes (Widen & Russell, 2010b). The results also vary according to the presentation mode, for example, texts versus pictures, and according to the emotion being expressed (Widen & Russell, 2010a). As regards the recognition of facial expressions, joy can be labeled correctly by the age of two (Widen & Russell, 2008). By the age of 4 or 5, all other basic emotions (Ekman & Friesen, 1975) are recognized (Camras & Allison, 1985). Pride is one of the non-basic emotions that are also recognized by 4-year-olds (Tracy, Robins, & Lagattuta, 2005). Differentiating emotions in stories is more difficult than in pictures. Although even 3-year-

olds can differentiate situations that elicit positive emotions from those that lead to negative emotions (Stein & Levine, 1989), it is more difficult to establish the causes of negative emotions. By the age of 5, children can distinguish fear and sadness (Harris, Johnson, Hutton, Andrews, & Cooke, 1989), whereas the discrimination between sadness and anger is difficult even for 6- to 7-year-olds (Levine, 1995).

Thus, research on the development of emotional knowledge suggests that even preschool children construct emotional inferences when they are asked to (Gnepp & Gould, 1985; Marmolejo & Heredia, 2006). Unfortunately, these data do not answer the question of the point at which children infer the character's emotional state during text comprehension, since only offline measures have been applied. It is possible that the children participating in those studies only thought about the character's emotion because they were asked a specific question about it. Thus, these results do not tell us whether children generate emotional inferences under realistic conditions.

Another question that has not been examined so far concerns possible differences between media in eliciting emotional inferences. Whereas most of the research on emotional inferences concentrates on written text, there is also evidence for emotional inference among adults watching films (Nieding & Ohler, 2004; Ohler et al., 2002). To our knowledge, no studies exist that examine emotional inference in auditory texts, but there are quite a lot of empirical data concerning the effects of different media on text comprehension.

Differences between Presentation Types

There are several lines of research that suggest there are differences between media concerning the elaboration of situation models. This is also relevant for research on emotional inferences since a superficial representation would not be sufficient to include information about the protagonist's emotion unless it is explicitly mentioned in the text. Given the perceptual and analogous nature of film, the perceptual symbol theories assume that narrative-

based films are more conducive to the construction of situation models than narrative-based written texts, because narrative films resemble everyday experience more closely. Supporters of this approach claim that the representations in the situation model are perceptual rather than amodal (Glenberg, 1997; MacWhinney, 1999). There is evidence that comprehenders mentally simulate perceptible as well as non-perceptible information, such as emotions or the sense of time (Barsalou, 2008; Kelter, Kaup, & Claus, 2004; Matlock, 2004; Zwaan, Madden, Yaxley, & Aveyard, 2004). Some results suggest that perceptible information enhances emotional inferences in comparison to non-perceptible information. For example, Gillioz, Gygax, and Tapiero (2012) found that behavioral information about emotion led to greater differences in reading times than emotional labels (e.g., "She danced all night" vs. "Suzanne was feeling happy", p. 2), which supports the claim that mental models and the included emotional inferences are of a perceptual nature. According to these results, the medium of film should be able to generate perceptually enriched situation models, because films integrate linguistic and perceptual symbol systems (Barsalou, 1999). As a result, situation models triggered by films should be able to generate inferences more easily. This is also supported by research on the modality effect, which shows that a dual presentation (audio and visual) is superior to a visual presentation (Leahy & Sweller, 2011; Mousavi, Low, & Sweller, 1995; Segers, Verhoeven, & Hulstijn-Hendrikse, 2008).

Another group of theories focus on the strain put on working memory by different presentation modes. These theories explain why particular media are better at imparting information at different ages. Baddeley and Hitch (1974) first described the tripartite structure of working memory. According to this model, working memory consists of the central executive and two subsystems. While the central executive functions as a master unit that monitors resources and information from different channels, the two subsystems are specialized on verbal (phonological loop) and visual-spatial information (visual-spatial sketchpad). Although the structure and the functional relationship of the three components of

working memory are mostly invariable throughout childhood (Michalczyk, Malstädt, Worgt, Könen, & Hasselhorn, 2013), the success in processing verbal or visual information is subject to developmental differences. In children up to the age of 5 years, the visual-spatial sketchpad is better developed than the phonological loop, so that younger children are more successful in processing visual-spatial information (Pickering & Gathercole, 2001). The development of visual-spatial processing continues as the children get older (Koppenol-Gonzalez, Bouwmeester, & Vermunt, 2012), while at the same time, between the age of 6 and 9 years, they begin making more use of the phonological loop (Hitch, Halliday, Dodd, & Littler, 1989), until they are eventually able to combine verbal and visual processing. These differences between age groups explain why the effects of different presentation modes are also subject to developmental processes. Several studies have shown that young children benefit from an audiovisual presentation form (Constantinidou, Danos, Nelson, & Baker, 2011; Gibbons, Anderson, Smith, Field, & Fischer, 1986; Pezdek, Simon, Stoeckert, & Kiely, 1987; Stoneman & Brody, 1983; Subrahmanyam, 2008). For instance, in the study by Gibbons et al., 4-year-olds benefitted from audiovisual presentations, in that they remembered more of the audiovisually presented information compared with the audio version, whereas no difference was found among older participants.

Since the phonological loop and the visual-spatial sketchpad are usually equally well developed in adults, they should not benefit from one or the other format. However, there are ambiguous results about this subject. For instance, Palmer (2000) found that adults prefer auditory coding, while Tibus, Heier, and Schwan (2013) found that adults process both types of information equally well.

Individual differences in background knowledge

Background knowledge about emotions was shown to enhance adults' ability to generate inferences in previous research (Gernsbacher & Robertson, 1992). We assume that this is true

for children as well. However, there are no studies so far that examine this assumption. Another factor that might be relevant for children's emotional inferences is their understanding of the symbol systems of various media. Every kind of medium is based on a certain set of symbol systems. Understanding these is essential for understanding the message. Films use a large variety of symbol systems, such as techniques like montage, sound effects, and camera parameters. Audiobooks can also use sound effects and, another important technique, the variation of voices to make clear which character is speaking.

Ohler (1994) assumes that the viewer's knowledge about a medium's symbol systems influences his or her construction of situation models. An example is a close-up of a protagonist, which can serve to emphasize the significance of this person (Ohler, 1994). Potter (1998) distinguishes between rudimentary and advanced skills of media literacy. Rudimentary skills are developed between 3 and 5 years (Potter, 2011). They include the fundamental abilities to recognize the symbols used by media, recognize patterns composed by these symbols and ascribe meaning to them. Nieding and Ohler (2008) call this "media sign literacy" ("Mediale Zeichenkompetenz"; p. 382). The results of our work group (Munk, Diergarten, Nieding, Ohler, & Schneider, 2012a; Munk et al., 2012b) suggest that the ages between 4 and 8 are crucial for the development of rudimentary media literacy skills. Advanced media literacy skills develop during school years and adulthood. They comprise analyzing, evaluating, abstracting, and appreciating a medium's message.

Purpose and design of the Present Research

The aim of the studies presented in this paper was to examine children's and adults' ability to infer the protagonist's emotional state in media. Specifically, our goal was to find out whether they generate these inferences online, that is, during the comprehension process. This is an important question, since the opportunity to learn about emotions from media can mean a great enhancement of their emotional knowledge. While understanding the emotion of a

protagonist in a movie, audiobook or written text, the child can learn about the emotions arising in various scenarios without having to experience all these situations him- or herself (Dorr, 1998). The effectiveness of acquiring emotional knowledge through media usage has not yet been evaluated, and the studies presented in this paper focus on an important foundation for this assumed educative effect.

Since we were interested in examining preschoolers as well as older children and adults, we chose two types of media that do not require reading skills. Therefore, the stories were presented as films (Study 1) or as audiobooks (Study 2). Both of these media are particularly popular with and frequently used by young children (Feierabend & Klingler, 2008; Grüninger & Lindemann, 2000; Rideout & Hamel, 2006). For the two studies, we designed an online measurement of emotional inferences for children. Reaction time (RT) measures have proven to be a suitable online inference measurement for children (Unsöld, 2008). Because our youngest participants had not yet learnt how to read, the commonly used method of reading latencies was not applicable to our studies. We applied the same logic as the reading latency measure, asking participants to react to a sentence that described an emotion. As with reading latencies, shorter RTs to sentences, where the emotion was the same as that of the protagonist, were judged to be an indicator of inference generation: if the participant built a situation model and included the character's emotion in this model, reacting to this same emotion should be facilitated. In order to judge the specificity of the inference, we differentiated between inconsistent emotions of both the same valence and the different valence: differences between consistent emotions and inconsistent emotions with the opposite valence were judged to be indicators of valence-related inferences, whereas differences between consistent emotions and inconsistent emotions with the same valence counted as exact inferences.

In addition, we measured the participants' emotional knowledge and media literacy in order to investigate the impact of these factors on the ability to generate emotional inferences.

Experiment 1

In our first experiment children and adults aged between 5 and 22 years watched short narrative films which were interrupted by spoken sentences that named an emotion of the protagonist (for example, “Timmy feels happy”). The emotions were either consistent (in the case of the given example “happy”) or inconsistent with the emotion of the protagonist induced by the storyline. The inconsistent emotions either had the same valence (in the case of the given example “proud”) or the opposite valence (in the case of the given example “sad” or “anxious”). The participants were asked to indicate whether the emotion stated in the sentence felt “nice” or “not nice” by pressing buttons and the reaction times (RTs) were recorded. In addition, the emotional knowledge and media literacy of the participants were measured.

Predictions

So far, no studies exist that measure children’s emotional inference generation during text comprehension online. Research on situation models has shown however, that 5-year-old children's situation models include the mental state of the protagonist (O'Neill & Shultis, 2007). In addition, there is empirical evidence that even children that young generate different types of inferences online (see Unsöld & Nieding, 2009). This leads us to the conclusion that 5-year-olds are also able to build emotional inferences online. Therefore our hypotheses are:

H 1: Online inferences on the emotional valence are constructed in every age group, i.e. RTs are shorter in the congruent than in the incongruent- opposite valence condition.

We also assume that the ability to emotionally infer in an online task develops with age, and especially that online emotional inferences become more precise as children get older, since the differentiation between emotions of the same valence is difficult for children under the age of 7 (Harter & Buddin, 1987).

H 2: 5-year-olds do not generate online inferences on exact emotional states. This ability develops with age.

We expect that both emotional knowledge and media literacy support emotional inferences, because higher emotional knowledge should help participants to differentiate between emotions of the same valence, and expertise in media literacy should be a foundation for building a more elaborate situation model (Ohler, 1994).

H 3: Participants with higher abilities in emotional knowledge construct more precise inferences than those with low abilities.

H 4: Participants with higher scores on the media literacy tasks construct more precise inferences than those with low scores.

Method

Participants. Sixty-six male and 66 female participants took part in this study. Four of them were excluded owing to missing data which left 128 participants, equally divided into four age groups each consisting of 32 participants (5-year-olds: 15 male, 17 female; 8-year-olds: 16 male, 16 female; 10-year-olds: 18 male, 14 female; adults: 15 male, 17 female). The mean age of the four groups was 5;5 (years; months) (SD = 0;3), 8;6 (SD = 0;4) , 10;8 (SD = 0;4), and 22;0 (SD = 3;2) respectively. The children were recruited from (pre-)schools in Würzburg, Germany; the adults were university students majoring in psychology.

Material

Emotional inferences. Twenty scenes from the *Lassie* TV series (1954-74, first shown in Germany in 1958) with an average length of 1.58 minutes were used as stimuli. The main protagonist of the TV series is a boy named Timmy, who lives with his parents and his clever collie Lassie on a farm. They experience many adventures in which Lassie saves people from dangerous situations. The series was not known to the children, but was to some of the adults. Only those scenes were chosen in which the main character Timmy felt joy, pride, fear or

sadness or was emotionally neutral. We chose these emotions because they are commonly recognized and understood by preschoolers (Tracy et al., 2005; Widen & Russell, 2010b), even though the emotions with the same valence might be confused by the youngest of our participants, as assumed in Hypothesis 2 (Harris et al., 1989; Janke, 2002). The neutral scenes served as a baseline and were not included in the results.

Which emotion arose in Timmy in which scene was decided according to Roseman, Antoniou, and Jose's (1996) model of emotional appraisal. We conducted a pilot study asking participants to name the emotion Timmy was feeling in each scene and how comprehensible it was judged on a four-point scale ranging from "not at all comprehensible" to "very comprehensible." Depending on the results we chose the four most appropriate scenes for each of the five different emotional states in terms of the unambiguousness of the particular emotional state and the comprehensibility of the scene's content. The scenes ended immediately after Timmy's emotional reaction; for example, after a close-up of his smiling face. After each scene a male voice said "Timmy feels …," accompanied by one of the four emotions happy, proud, sad or anxious.[1] This emotion either matched or mismatched the emotion that Timmy felt in the scene. The mismatching emotions were either of the same or of the opposite valence of Timmy's actual emotion. Thus, three conditions of congruency resulted from this variation. For example, in one episode, Timmy was given the pony he wished for as a present, and therefore he felt happy. In the congruent (CON) condition, the spoken sentence was "Timmy feels happy"; in the incongruent with same valence (INCON-SAME V) condition, the sentence was "Timmy feels proud"; and in the incongruent condition with opposite valence (INCON-OPP V) it was either "Timmy feels sad" or "Timmy feels anxious."

[1] The film scenes and sentences were presented in German, since all the participants were German. The German sentences were "*Timmy fühlt Freude*", "*Timmy fühlt Stolz*", "*Timmy fühlt Trauer*", and "*Timmy fühlt Angst*".

Directly after a single scene was presented, a blue dot appeared on the screen and a male voice uttering one of the "Timmy feels…" sentences was presented. The participants were asked to press one of two keys as quickly as possible, depending on the emotion stated by the male voice: the "smiley" key (☺) for emotions that felt nice and the "frowny" key (☹) for emotions that did not feel nice. The reaction time and whether the response was correct or false were recorded.

Questionnaire on general understanding of the film scenes. In addition, we designed a questionnaire containing one simple question for each scene. These questions were very simple to answer and served the purpose of guiding participants' attention to the film. They also gave us the opportunity to judge and compare participants' understanding of the content of the film clips. These questions did not focus on Timmy's emotion and therefore did not serve as an indicator of emotional inferences. The latter were judged by the RTs only. The questions were asked by the experimenter after each scene, after the participant reacted to the "Timmy feels…" sentence by pressing one of the two keys.

Emotional knowledge. As a measurement of emotional knowledge Janke's German version (2008) of Pons, Harris, and de Rosnay's emotion comprehension test (2004) was used to test the children. Eighteen items measure nine different aspects of emotional knowledge. The test assesses understanding of external, mental, and reflective aspects of emotion. Examples of the scales on external signs are situational knowledge and facial expressions. The questions on mental indicators focus on false belief. The reflective elements include knowledge about multiple emotions. For an overview of the nine subscales, see Pons et al. (2004). For each question, four pictures of different emotional states were depicted. The child indicated his or her answer by pointing to the one that he or she deemed correct. The scoring was carried out according to the original test (Pons et al., 2004): for each of the nine subscales 1 point was scored if the child's correct answers exceeded a certain score; for the first two components (Recognition and External Cause), four of the five items had to be answered

correctly to gain the point. For the other seven components, all answers had to be correct in order to gain the point. The total score could thus vary from 0 to 9. This total score was again divided by 9 for a standardized score varying from 0 to 1. This was necessary to account for the different measures for different age groups in order to gain a comparable score for each test.

Most of the aspects of emotional knowledge measured in this test are based on theory of mind, except for the scale on facial emotion recognition. Since healthy individuals have developed a firm understanding of the different components of theory of mind during their preschool and elementary school years (Flavell, Green, & Flavell, 2000; Miller, 2009; Talwar & Lee, 2008; Wellman, Cross, & Watson, 2001), the test is reliable only for children aged 3 to 11. For this reason we gave adults a test containing emotion recognition only. Seventy pictures from Ekman and Friesen's (1976) collection were presented on a computer screen. The depicted emotions were joy, surprise, sadness, anger, disgust, and fear. Each emotion was presented ten times. The remaining ten pictures showed neutral expressions. Pictures were presented in a fixed order that made sure that the same emotion was not presented more than twice in a row. For each picture participants were asked to choose the appropriate answer from the above list of seven options. The maximum score for this measure was 70 points. The standardized score varied from 0 to 1.

Media literacy. We developed questionnaires for every age group which measured the seven dimensions of media literacy described by Groeben (2002). There were several questions for each of these subscales, most of them presented in a forced-choice format and some posed as open questions. Different questions were used for children and adults in order to provide an appropriate level of difficulty, which was determined by means of a pilot study. Appendix A gives an overview of the seven subscales, providing examples for each of them. The result for each of the seven subscales was divided by its maximum available score,

resulting in a maximum score of 1 point for each subscale. The sum of these subscores formed the total score with a maximum of 7 points (the standardized score varied from 0 to 1). Children completed the questionnaire with the help of the experimenter, who read the questions aloud and wrote down their answers. Adults read and answered the questions themselves.

The aforementioned questionnaires focused on advanced skills. Since the preschoolers' abilities can be better described by the measurement of rudimentary skills, a test for media sign literacy developed by Domaratius and Ohler (2006) was used. This online measure was conducted on a computer. An animated tutor in the character of a monkey introduced the children to the test and explained the different tasks, which were presented in a forced-choice format. The children chose an answer with the computer mouse. The tutor praised them for their efforts but did not comment on the correctness of their choices. The test measured media sign literacy in 11 subscales which are listed in Appendix B. For each subscale, the total score was divided by the maximum available score, resulting in a maximum of 1 point for each scale. The sum of those 11 subscores formed the total score with a maximum of 11 points. Divided by 11 they provided a standardized score varying from 0 to 1. Only the preschool children completed this measure, since a pilot study had shown that it was too easy for older children and adults.

Design and Procedure

Each participant was tested individually in a quiet room provided by the kindergarten, school or university. In the first session, the preschoolers first completed the online media literacy test, followed by the media literacy questionnaire. The second session comprised the emotion comprehension test, the third session the test on emotional inferences. The schoolchildren had two sessions each: the first comprised the media literacy questionnaire and the emotion comprehension test, the second the emotional inference test. Adults were tested in a single

session. They completed the tests in the same order as the children. The scores of the pretests were calculated by the experimenter immediately after the test. Every age group was split into two groups according to the group's median in each test.

The experimental video scenes were randomly assigned and presented on a computer screen, using the program Presentation®. There were four conditions that counterbalanced episodes and condition. This resulted in a 3 (congruency: congruent [CON], incongruent – same valence [INCON-SAME V], incongruent – opposite valence [INCON-OPP V]) x 4 (age) design, with congruency as within-subjects variable. The participants were assigned to one of the four conditions. We aimed to assign an equal number of participants with pretest scores beneath and above the median and an equal number of female and male participants to each condition. Aside from that the assignment was random. Each participant saw 20 experimental scenes (four consistent, four inconsistent – same valence, eight inconsistent – opposite valence, four neutral), requiring ten "smiley" responses and ten "frowny" responses.

Data Analysis. In a first step, all false reaction times and extreme RT values were excluded from further analysis. We defined a criterion of three times the standard deviation higher or lower than the mean of the respective age group for detecting extremes (see Rey, 2012). We did not substitute any missing values, because as well as the question of whether inferences were generated at all we were also interested in individual differences in the ability to generate inferences; therefore substitution did not seem appropriate. This had a disadvantage in that four participants had to be eliminated because too many of their data (more than five responses) were missing after this procedure. In a second step, we conducted ANOVAS to test for effects of congruency, age, emotional knowledge, and media literacy. If the ANOVA's main effect was significant, *t*-tests compared the RTs further. In the case of emotional knowledge and media literacy, we calculated further *t*-tests if the ANOVA showed a significant interaction effect between these competencies and emotional inferences. We divided the participants of each group into two groups with high or low competencies using

the median of the respective scores as a cut-off point. Since the participants of the four different age groups completed different versions of the tests (in the case of media literacy) or entirely different measures (emotional knowledge), their scores on the various measures varied not only because of their skills but also because of differences between the measurements. This meant we simply entered the information into the *t*-test on whether the individual participants were below or above the median in order to avoid variance because of the different measures, in an attempt to make the results of the *t*-tests of the different age groups more comparable.

Results and Discussion

Results. The mean number of errors in the reaction time task was very low in every age group (5-year-olds: 1.56% [SD = 3.22], 8-year-olds: 0.78% [SD = 2.24], 10-year-olds and adults: 0.31% [SD = 1.23] in both groups). These RTs were deleted. Extreme values were also deleted before further analysis. This applied to 4.37 % [SD = 7.49] of the 5-year-olds' RT values and 2.34 % [SD = 4.40], 2.19 % [SD = 3.10], and 2.03 % [SD = 4.37] for the three older age groups respectively.

The scores in the questions on general understanding on the film were quite high: the 5-year-olds answered 91% of the questions correctly (SD = 6.32), the 8-year-olds 99% (SD = 2.58), the 10-year-olds 98% (SD = 2.55), and the adults 97% (SD = 3.39). The low error ratio and the high score in the questionnaire imply that the participants had no difficulties following the storyline or in mastering the RT task.

The mean standardized scores in the tests of emotional knowledge were 0.50 (SD = 0.14), 0.77 (SD = 0.16), and 0.94 (SD = 0.08) for 5-, 8-, and 10-year-olds respectively. The age difference was significant (F (2,93) = 94.12, $p < 0.001$, $\eta_p^2 = 0.67$). The adults' mean score in the emotion recognition test was 0.85 (SD = 0.07). With regard to media literacy, the mean score of the 5-year-olds in the online measure was 0.68 (SD = 0.08). The mean scores in

the questionnaire were 0.41 (SD = 0.11), 0.51 (SD = 0.10), 0.62 (SD = 0.08), and 0.57 (SD = 0.12) for the four age groups. Since there were different versions of the test for preschoolers, schoolchildren, and adults, we cannot compare these values between the age groups.

A four-factorial ANOVA tested the effects of congruency, age, emotional knowledge, and media literacy on the RTs. We entered the standardized scores of the emotional knowledge and media literacy tests (only the questionnaires, not the online measure) into the ANOVA. Since the neutral scenes (N) served as distractors, those RTs were not included in the statistics. The main factor of congruency was significant, F (2,244) = 4.11, $p < 0.05$, η_p^2 = 0.03, and so was the main effect of age, F (3,122) = 23.06, $p < 0.001$, $\eta_p^2 = 0.36$. All interactions with congruency were significant or marginally significant (by age: F [6,244] = 5.02, $p < 0.001$, $\eta_p^2 = 0.11$; by emotional knowledge: F [2,244] = 2.89, $p = 0.058$, $\eta_p^2 = 0.02$; by media literacy: F [2,244] = 4.15, $p < 0.05$, $\eta_p^2 = 0.03$). Given the interaction effects, the main effect will not be interpreted at this point.

***The role of age*.** Figure 1 displays the RTs of the four age groups in the three congruency conditions.

- Enter Figure 1 about here -

The *t*-tests showed decreasing RTs in the older age groups. The difference between 5-year-olds and every other older age group was significant (t [62] = 6.05, t [62] = 6.97, and t [62] = 10.15 for the differences with 8- and 10-year olds and adults respectively, $p < 0.001$ each), which is because of the longer RTs in the younger age groups. The difference between 8- and 10-year-olds was not significant (t [62] = 1.77, $p = 0.08$), but both 8- and 10-year-olds showed significant differences compared with adults (t [62] = 6.98 and t [62] = 4.52, $p < 0.001$ each), also due to children's longer RTs.

The main effect of congruency in the ANOVA might have been distorted by the interaction effects. In order to find out what kind of inference (exact inference: difference between CON and INCON-SAME V vs. valence inference: difference between CON and INCON-OPP V) was generated in which age group, *t*-tests comparing the congruency conditions were repeated for each group. For both 5- and 8-year-olds the difference between CON and INCON-SAME V was not significant (5-year-olds: t [31] = -0.65, p = 0.52; 8-year-olds: t [31] = -1.12, p = 0.27), but the difference between CON and INCON-OPP V produced significant results (5-year-olds: t [31] = -4.17, $p < 0.001$; 8-year-olds: t [31] = -5.37, $p < 0.001$). This shows that younger children do generate emotional inferences, but only include the emotional valence of the character's emotion and not the exact emotional state. Ten-year-olds and adults both showed RT differences indicating exact inferences (10-year-olds: t [31] = -2.43, $p < 0.05$; adults: t [31] = -2.14, $p < 0.05$) and inferences of emotional valence (10-year-olds: t [31] = -4.22, $p < 0.001$; adults: t [31] = -3.40, $p < 0.01$). Our first and second hypotheses are thus supported: inferences are generated in every age group, but constructing exact emotional inferences seems to be an ability that develops with age.

***The role of emotional knowledge and media literacy*.** As reported above, the interaction between congruency and emotional knowledge and the interaction between congruency and media literacy were both significant. To confirm these results for the four age groups, the *t*-tests calculated above (comparing the RTs of every age group in CON, INCON-SAME V and INCON-DIFF V) were calculated again for individuals with high or low emotional knowledge. Using the median score of emotional knowledge we built two subgroups in each age group: those participants with scores above the median constituted the group with "high emotional knowledge," and those with scores below the median formed the group with "low emotional knowledge." The *t*-tests were calculated separately for each group (see Table 1 for the results of the various *t*-tests).

- Enter Table 1 about here -

The results showed that high emotional knowledge was associated with emotional inference in every age group except for the 8-year-olds. For 5-year-olds and adults no inferences of any kind were shown for the groups of low emotional knowledge, but inferences of emotional valence were shown for the groups of high knowledge. Ten-year-olds constructed emotional inferences of valence in both groups, but exact inferences were only evident for those with high emotional knowledge. Emotional knowledge did not seem to have an effect on 8-year-olds' ability to construct emotional inferences, since the same results were detected for groups of high and low knowledge: both generate inferences based on emotional valence, whereas no effect resulted for exact inferences. Hypothesis 3 was affirmed for every age group except the 8-year-olds.

We used the same median split to divide each age group into two subgroups of higher and lower media literacy. The results of the *t*-tests comparing emotional inferences for each of these groups are depicted in Table 2.

- Enter Table 2 about here -

For 8- and 10-year-olds and adults the same result patterns for groups with high and low competency levels that had already been found with emotional knowledge as a cut-off point were detected; media literacy had no effect on 8-year-olds' ability to construct emotional inferences, whereas for 10-year-olds and adults the ability was enhanced as described above. Hypothesis 4 was supported for every age group except the 8-year-olds. The results for the 5-year-olds will now be reported in more detail, because the two measures of media literacy interestingly led to different results; using the results of the questionnaire as a cut-off point, we found only those children with high scores generated inferences at all, and only inferences based on the emotional valences. Using the outcomes of the online computer measure, again we found only those with high scores emotionally inferred, not only about the emotional

valence but also about the exact emotional state. Therefore, for the 5-year-olds the computer measure seems to be associated with emotional inference more closely than the questionnaire.

Discussion. Experiment 1 showed that children from the age of 5 generated inferences about a movie character's emotional state. While 5- and 8-year-olds only generate inferences of the emotional valence, 10-year-olds and adults infer the character's exact emotional state. After splitting up the groups into participants with media literacy or emotional knowledge above or below the median, we found that only the 5-year-olds in the groups above the median inferred the emotional state, while the children in the groups below did not generate any inferences. The children in the group above the median of media sign literacy were even able to generate exact inferences. A similar pattern of results was displayed for the adults: only those in the above-the-median-group generated inferences. Exact emotional inferences were not significant, but the trend was the same. This result was unexpected, since we found both kinds of inferences for the adults before splitting up the groups at the median. We assume that this result is due to the fact that the RT-task was fairly simple for the adults, which resulted in small differences between the congruency conditions. This and the reduced power in the *t*-tests with only half of the participants might have led to the non-significant result.

The 8-year-olds built emotional inferences only for emotional valence. Neither emotional knowledge nor media literacy was associated with their inference skills. Since the same pattern of 8-year-olds' inference was found in Experiment 2, we will come back to this matter in the general discussion. The 10-year-old participants were generally capable of building emotional inferences. Additional competencies in either emotional knowledge or media literacy were associated with the ability to build inferences not only about the valence of the character's emotion, but about the exact type of emotion he or she experienced.

Experiment 2

Because of the encouraging results of Experiment 1 we went on to examine the question of whether children are able to build emotional inferences not only in audiovisual texts, like movies, but also for auditory presentations like audiobooks. We attempted to design the second experiment to be as similar to Experiment 1 as possible. Therefore we used the movie scenes of Experiment 1 as a model for creating audiobooks.

Predictions

So far no studies exist that test whether emotional inferences are built during the reception of an audiobook. Since all the necessary information for understanding a character's emotion is given by the narrator (see Method section of Experiment 2), it seems likely that the audio version will promote emotional inferences just as much as the audiovisual version. We also expected that emotional knowledge and media literacy would be similarly important in auditory texts. Therefore our hypotheses H1 to H4 are the same as for Experiment 1.

Method

Participants. A total of 145 participants took part in this study. Seventeen of them were excluded owing to disruptions during testing (7) or missing data (10), which left 128 participants, equally divided into four age groups, each consisting of 32 participants (5-year-olds: 15 male, 17 female; 8-year-olds: 16 male, 16 female; 10-year-olds: 16 male, 16 female; adults: 15 male, 17 female). The mean age of the four groups was 5;7 (years; months) (SD = 0;3), 8;1 (SD = 0;4), 10;2 (SD = 0;4), and 22;4 (SD = 3;6) respectively. The children were recruited from (pre-)schools in Würzburg, Germany; the adults were university students majoring in psychology. None of the participants had previously taken part in Experiment 1.

Material. Based on the 20 movie scenes used in Experiment 1, scripts for 20 audiobook sequences were written and recorded by a male narrator. The narrator reported

everything that was happening in an enthusiastic manner. In between his comments the original voices of the actors and the music soundtrack of the movie scenes were recorded. In that way we ensured that the participants received all the necessary information that the participants of Experiment 1 had seen, while at the same time keeping it as similar to the movie scenes as possible by using the original soundtracks. The resulting audiobook sequences, similarly to the movie scenes, had an average length of 2.02 minutes. We decided to include a screenshot of Timmy's facial expression at the end of each sequence for the following reasons: participants of Study 1 were able to see Timmy's face throughout the story. We cannot tell from the results whether the participants inferred Timmy's emotional state only by reading his facial expression. We reasoned that if we presented the audiobooks without pictures and did not find any emotional inferences, we would not be able to tell whether this was due to the inability to read the facial expression or to differences between the two media. Thus, during the last 5 seconds of each audiobook sequence a screenshot of the respective movie scene was shown. This screenshot was chosen from the last 5 seconds of the movie scene and usually contained a close-up of Timmy's face, showing his emotional reaction. If Timmy's face was not on screen in the last 5 seconds, a different screenshot was chosen from this time span.

The ensuing procedure to measure emotional inferences was identical to Experiment 1, with the same male voice saying "Timmy feels ...," accompanied by one of the four emotions happy, proud, sad or anxious, and the identical recording of these sentences from Study 1. Also, for the questions for general understanding, emotional knowledge, and media literacy we used the same measures as in Experiment 1.

Design and Procedure

The procedure was identical to Experiment 1. Each participant was tested individually in one, two or three sessions (adults vs. schoolchildren vs. preschoolers) in a quiet room provided by

the kindergarten, school or university. The order of the tests was the same as in Experiment 1. Again, the median of the pretest scores was assessed and we took account of this median in assigning participants to the four conditions in the emotional inference test.

The audiobooks were presented on a computer, using the program Presentation®. The participants were asked to look at the blue dot in the middle of the screen while listening to the audiobooks. The screenshots taken from the movie scenes were presented at the same spot, substituting the dot at the end of each audiobook. They were shown for 5 seconds, during the last 5 seconds of the audiobooks. Directly afterwards the screen went back to showing the blue dot and the recording of the male voice saying the target sentence was being played. As in Experiment 1, participants were asked to press one of the two keys as fast as possible after hearing the male voice saying the sentence. The same four conditions counterbalancing episodes and congruency conditions were used as in Experiment 1.

For data analysis we used the same methods as described in Experiment 1.

Results and Discussion

Results. As in Experiment 1, RTs for incorrect responses and outliers in the reaction time task were excluded from further analysis. The mean number of errors was very low in every age group (5-year-olds: 2.03% [SD = 3.99], 8-year-olds: 0.94% [SD = 2.36], 10-year-olds: 0.47% [SD = 1.48], adults: 0.63% [SD = 2.11]). The percentage of outlier values was 5 % [SD = 8.96] for the 5-year-olds and 2.19 % [SD = 5.07], 3.24 % [SD = 7.06], 2.78 % [SD = 6.81] for the three older age groups respectively. Also similar to Experiment 1 were the high rates of correctly answered questions on general understanding of the audiobooks: 5-year-olds: 87% (SD = 8.28), 8-year-olds: 96% (SD = 5.77), 10-year-olds: 97% (SD = 4.39), adults: 99% (SD = 2.14). These rates imply that the participants did not struggle to follow the storyline or accomplish the RT task.

The mean standardized scores in the test of emotion comprehension were 0.57 (SD = 0.16), 0.82 (SD = 0.12), and 0.88 (SD = 0.09) in the three children's groups. The score rose significantly with age (F (2,93) = 54.11, $p < 0.001$, $\eta_p^2 = 0.54$). The adults' score was 0.86 (SD = 0.06). In the media literacy online test, the preschoolers' mean score was 0.72 (SD = 0.07). The mean scores in the questionnaire were 0.43 (SD = 0.09), 0.53 (SD = 0.09), 0.59 (SD = 0.08), and 0.60 (SD = 0.07) for the four age groups. As in Experiment 1, we did not compare these results because of the different versions of the tests in the four age groups. A four-factorial ANOVA tested the effects of congruency, age, emotional knowledge, and media literacy. As in Experiment 1, we entered the standardized scores of emotional knowledge and media literacy because of the different ranges of the test. The main factor of congruency was not significant ($F < 1$), but the main effect of age (F [3,122] = 48.35, $p < 0.001$, $\eta_p^2 = 0.54$) was statistically significant. Only the interaction between congruency and age was significant (F [6,244] = 3.21, $p < 0.01$, $\eta_p^2 = 0.07$); those between emotional knowledge and media literacy were not (both: $F < 1$). Therefore no further analyses were conducted concerning these abilities. Hypotheses 3 and 4 are rejected.

The role of age. Figure 2 displays the RTs of the four age groups in the three congruency conditions in Experiment 2.

- Enter Figure 2 about here –

The *t*- tests showed decreasing RTs in the older age groups. The differences between each age group and all older age groups were significant (5-year-olds to every older age group: t [62] = 6.08, t [62] = 9.78 and t [62] = 16.21, $p < 0.001$ each; 8-year-olds to 10-year-olds: t [62] = 2.92, $p < 0.01$; 8- and 10-year-olds to adults: t [62] = 9.75 and t [62] = 9.19, $p < 0.001$ each). All effects were the result of longer RTs in the younger age groups.

As stated above, the main effect of congruency of the calculated ANOVA was not significant, but this might be due to the interaction with age. In order to find out whether

inferences (exact vs. valence) were generated in any of the four age groups, *t*-tests were calculated for each group. For both 5- and 8-year-olds the difference between CON and INCON-SAME V was not significant (5-year-olds: t [31] = -0.28, p = 0.78; 8-year-olds: t [31] = -1.00, p = 0.33), but the difference between CON and INCON-OPP V lead to significant results (5-year-olds: t [31] = -2.75, p = 0.01; 8-year-olds: t [31] = -5.64, p < 0.001). Ten-year-olds showed both exact inferences (t [31] = -2.04, p = 0.05) and inferences of emotional valence (t [31] = -5.53, p < 0.001), adults only exact inferences (t [31] = -3.73, p = 0.001) but no inferences of emotional valence (t [31] = -0.78, p = 0.44).

Discussion. As in Experiment 1, our first two hypotheses can be supported: participants of all age groups built emotional inferences, but only 10-year-olds and adults could infer the exact emotion of the protagonist, whereas the younger children only inferred the valence of the character's emotional state. In the case of the adults, only those RT differences pointing to exact emotional inferences were significant, and not those for inferences on emotional valence. This cannot mean that they did not infer the emotional valence, because obviously this is included in the exact emotional inference. Instead, we think that the "novel-popout" effect (Johnston & Hawley, 1994, p. 56) might be responsible for this result. The "novel-popout" effect describes the phenomenon whereby novel cues are easily recognized from a crowd of familiar distractors. Accordingly, the adults' RT differences can be explained by the claim that they integrated the emotion described in the audiobook into their situation model, so it became familiar to them. If an emotion with an opposite valence was presented, it was so different from the one in the situation model that it popped out and thereby enabled a fast response. Those with the same valence were not different enough to cause a popout effect and therefore took longer than those with the opposite valence. This is also in line with Treisman and Gelade's feature-integration theory (1980), which claims that the more similar targets and distractors are, the longer the processing time will be. An explanation for the lack of this effect in Study 1 will be given in the General Discussion.

With the exception of this latter finding, the results of Studies 1 and 2 were quite similar. Presentation mode did not seem to have a high impact on emotional inference. There might have been an effect on RTs, however, or on how well participants understood the stories, as analyzed in the following section.

Comparison of Study 1 and Study 2

To examine the influence of the presentation mode, the scores of the questionnaire on general understanding were compared in a two-factorial ANOVA with presentation mode as a between-subjects factor. Since the literature cited in the Introduction suggests that the effect of presentation mode differs with age, this ANOVA was calculated separately for each age group. The main effect of the presentation mode was significant for the 8-year-olds and the adults (8-year-olds: F [1,62] = 7.83, $p < 0.01$, $\eta_p^2 = 0.11$; adults: F [1,62] = 4.87, $p < 0.05$, $\eta_p^2 = 0.07$; 5-year-olds: F [1,62] = 3.42, $p = 0.07$, $\eta_p^2 = 0.05$; 10-year-olds: F [1,62] = 2.03, $p = 0.16$, $\eta_p^2 = 0.03$). The *t*-tests showed that this was based on greater understanding of the audiovisual mode in the case of the 8-year-olds, whereas the adults scored higher in the auditory mode (percentage of correct answers: 8-year-olds: Study 1, audiovisual: $m = 98.86$, $SD = 2.58$; Study 2, auditory: $m = 95.74$, $SD = 5.77$; adults: Study 1, audiovisual: $m = 97.02$, $SD = 3.39$; Study 2, auditory: $m = 98.58$, $SD = 2.14$).

Similarly, the three-factorial ANOVA was conducted with regard to the mean RTs. The main effect of the presentation mode was only significant for the 8-year-olds (F [1,62] = 7.78, $p < 0.01$, $\eta_p^2 = 0.11$; 5-year-olds: F [1,62] = 1.79, $p = 0.19$, $\eta_p^2 = 0.03$; 10-year-olds: F [1,62] = 3.21, $p = 0.08$, $\eta_p^2 = 0.05$; adults: F [1,62] = 2.31, $p = 0.13$, $\eta_p^2 = 0.04$). The ensuing *t*-tests showed that the results for the 8-year-olds were due to shorter reaction times in Study 1, that is, the audiovisual presentation mode (Study 1, audiovisual: $m = 999$ ms, $SD = 124$; Study 2, auditory: $m = 1107$ ms, $SD = 179$).

To sum up, the 8-year-olds' understanding is advanced by the audiovisual presentation, which is supported both by shorter RTs and higher general understanding. Conversely, the adults' understanding seems to be enhanced by auditory presentation. However, although the differences concerning general understanding proved to be statistically significant, they were very small (8-year-olds: 99% of correct answers in the auditory version vs. 96% in the audiovisual version; adults: 99% vs. 97%). The questionnaires seem to have been too easy. While the 8-year-olds' preference for the audiovisual version finds further support by the difference in RTs, this is not the case for a possible preference of the adults concerning the auditory version. A more sensitive measure should be applied before further conclusions are possible concerning adults' preferences for presentation mode.

The results reported in this section need to be analyzed with caution for two more reasons: first of all, they are based on two different samples. Therefore the differences we found might be due to differences between the samples. Second of all, the picture presented at the end of the audio version makes the interpretation more difficult. We will discuss these limitations in the General Discussion.

General Discussion

The aim of our studies was to find out whether children and adults are able to understand the main character's emotional state of a movie or an audiobook online, that is during the viewing or listening process. The data supported our hypothesis that children between the age of 5 and 10 years and adults are able to construct emotional inferences, i.e. they mentally represent the emotions of a protagonist in the situation model, both when watching film clips and when listening to audiobooks. Also according to our hypotheses is the conclusion that the ability to construct emotional inferences depends on several factors. First of all, it seems to develop with age. Although even 5-year-olds constructed emotional inferences, they were only able to make inferences on the basis of the affective valence of the character's emotional state. As

children get older, though, their inferences become more precise, until they are able to construct exact emotional inferences by the age of 10. Individual background knowledge on emotion and media literacy were also connected with the ability to build more precise emotional inferences, but only in Experiment 1, that is, the audiovisual presentation mode. Aside from that, the presentation mode did not have any effects on children's emotional inferences and only a small effect on adults' inferences. This corresponds with the findings of Kendeou, Bohn-Gettler, White, and van den Broek (2008), which showed that children's inference generation skills were highly interrelated across aural, visual, and written stories.

Overall, our results support our assumption that even 5-year-old children are able to build profound situation models without being asked to do so, that is during their natural comprehension process. These findings are in line with research that has shown other types of online inferences in preschool children (e.g. Unsöld, 2008). Our results add to the current knowledge about children's situation models by demonstrating that inferences about the character's emotional state are integrated into those situation models. That young children build inferences based only on the affective valence and that their inferences get more precise with age is in agreement with Janke's (2002) findings that children under the age of 7 have difficulty differentiating between emotions of the same valence.

In the following paragraphs, we will discuss the influence of background knowledge and presentation mode on emotional inferences.

The Role of Media Literacy and Emotional Knowledge

In our first experiment with the audiovisual presentation, emotional knowledge was associated with constructing emotional inferences in every age group, except for 8-year-olds. Specifically, we found that 5-year-olds and adults showed emotional inferences only in the groups with high emotional knowledge. While this makes sense for the 5-year-olds, the adults' result might be due to the fact that the task was too simple for them, as explained in the

discussion of Experiment 1. Ten-year-olds showed inferences of emotional valence in both groups, but children with high emotional knowledge drew inferences on the basis of specific emotions. This finding matches those of Gernsbacher and Robertson (1992), i.e. that previous knowledge is important for the generation of emotional inferences.

With regard to media literacy, in Experiment 1 we found the same effect as that of emotional knowledge on 10-year-olds and adults and no effect on the 8-year-olds. Concerning the 5-year-olds, media literacy seemed to be more closely associated with constructing emotional inferences than emotional knowledge, since children with high scores in one of the media literacy measures were able to construct exact emotional inferences. This confirms our hypothesis that a person's media literacy modifies how he or she receives and processes films. Not surprisingly, for the 5-year-olds only the computer measure, not the questionnaire, led to this result. This is in accordance with Potter's (1998) assumption that advanced skills, which are measured by the questionnaire, begin to evolve at the end of primary school. In earlier childhood, though, primarily rudimentary skills are developed, and therefore a test assessing those skills is understandably more sensitive in differentiating the state of development in children of that age. Since we did not have the same tests for all age groups, the question about which test is better suited to measuring media literacy skills relevant for emotional inference can only be answered for the 5-year-olds, because only they completed both types of tests. Further research would be necessary to analyze whether a measure of rudimentary skills would have been more suitable for the older children and the adults.

Why did neither media literacy nor emotional knowledge have an influence on the inference generation of 8-year-olds? A close inspection of how inference processes seem to develop, in the light of this study, might help to explain these findings. 5-year-olds only drew emotional inferences when they had a firm foundation either in media literacy or in emotional knowledge. Ten-year-olds, on the other hand, constructed emotional inferences regarding the emotional valence, but only those who were good in either media literacy or emotional

knowledge built exact inferences. For the 5-year-olds, inference generation of emotional valence seems to be in their zone of proximal development (Vygotsky, 1978), i.e. the difference between what a child can do with and without help. The 10-year-olds are already good at constructing inferences of emotional valence, but building exact emotional inferences is within their reach. Media literacy and emotional knowledge both seem to be skills that are associated with children reaching their zone of proximal development concerning emotional inferences, because both general world knowledge and media literacy are important for the construction of situation models and inference generation (Gernsbacher & Robertson, 1992; Ohler, 1994). 8-year-olds are between the developmental stages of 5- and 10-year-olds. They were already drawing inferences about the emotional valence of a character's state, but could not match the abilities of the 10-year-olds. Therefore, neither of the assessed competencies was associated with any detectable change in their inference generation.

In Experiment 2, we did not find any influence of media literacy or emotional understanding on emotional inference. We think that this effect is not explainable by the presentation mode alone, but by the combination of the auditory presentation mode and the measurements for media literacy and emotional knowledge, which were the same as in Experiment 1. With regard to media literacy, the effect might be due to the fact that both the questionnaire and the computer test assess competencies concerning audiovisual media (TV and DVD) more than auditory media. Although there are some items in the questionnaire concerning print media, there are hardly any about audiobooks. So this questionnaire measures competent use of audiovisual media rather than media in general.

As regards emotional knowledge, our results might be due to a similar reason: our measurements for emotional knowledge were mainly based on pictures. This is most apparent in the adults' task of labeling Ekman's pictures, but it is also true for the children's task. Children gave their answers by pointing to the picture with the emotional facial expression that they thought was correct. Hence, for every scale of this task, which differed in the amount

of verbal information provided, it was always necessary to give a picture-based answer. Therefore the ability to recognize facial expressions was more important than eliciting emotional states from verbally given information. This explains why both the children's and the adults' measurements for emotional knowledge were more decisive for emotional inference in the audiovisual presentation mode than the auditory mode.

Differences between Presentation Modes

With regard to the differences between audiovisual and auditory presentation modes, we found little difference concerning emotional inferences, except for the group of adults. The difference between their RTs on CON and INCON-OPP V was not significant, but the difference between CON and INCON-SAME V was. This indicates exact emotional inferences that should include the valence of the emotion. We explained this finding by the "novel-popout" effect. The question remains why this effect only appeared in the second, not in the first, experiment. This might be due to the difference in presentation mode. However, this assumption is not supported by our results. As explained above, there was no difference between the auditory and the audiovisual group concerning reaction times. The difference concerning general understanding was statistically significant, but too small to be considered relevant (99% vs. 97%). Further studies with a more difficult measure for general understanding are necessary in order to explain these findings.

Although no difference was found concerning inference, the 8-year-old children seemed to benefit from audiovisual presentation. As with the adults, the difference in the understanding of story content was significant, but very small (99% vs. 96%), yet their preference for the audiovisual version finds further support by the difference in RTs.

No differences were found for the groups of 5- and 10-year-olds. A closer look at the longitudinal results of Palmer (2000) can help to explain this finding. She found that children up to the age of 5 did not show a preference, 6-year-old children preferred visual coding, 7-

year-olds audio-visual coding, and from the age of 8, children adopted the auditory preference of adults. Our results match these findings concerning the 5-year-olds, but the older children in our study made the shift to auditory preference later than those in Palmer's. This could be due to a need for deeper understanding in our task compared with the requirement to memorize pictures in Palmer's. It can also be explained by the findings of Koppenol-Gonzalez et al. (2012), who showed that while 6- to 9-year-olds do become more skilled in processing verbal information, it doesn't mean that it gets easier than processing audiovisual information, since children still become more experienced in this latter type of media as well.

Limitations and Conclusions

Since we used a picture of Timmy's face at the end of each audiobook, Study 2 was not an auditory-only condition. It is therefore not clear whether participants would have been able to infer Timmy's emotional state if it had not been for the picture included at the end of the audiobooks. Although we did gain insights into the processing of different media concerning situation model construction, we cannot tell from our results whether children and adults build emotional inferences in auditory-only conditions, that is, under realistic circumstances at home, where audio-books do not come with pictures. What we examined resembles more the condition where a picture-book is read to a child who can simultaneously hear his or her parent's voice and see the pictures in the book. Further research with an auditory-only condition is necessary for greater understanding of emotional inference in auditory texts. We are currently preparing two studies that compare audiovisual and auditory conditions in a within-subject design. We are also planning to vary these studies concerning the picture at the end to gain further insights into the role of this picture. Therefore, these planned studies will address another limitation of the current studies, since they will use a within-subject design. Although the groups of participants in the two studies were similar in terms of age, gender, and education, they were not randomly assigned to the audiovisual or auditory versions. We

can therefore not exclude the possibility of group effects. A within-subject design will also provide further possibilities for statistical analysis.

A third limitation concerns the fact that we used different measures for emotional knowledge and media literacy between the age groups, owing to the lack of measures suitable for both children and adults. This limits our conclusions concerning the differences in the influence of these skills in different age groups. These conclusions are further limited by the fact that we used correlational methods only and are therefore unable to make any causal interpretations. We can therefore not conclude that media literacy and emotional knowledge are the key factors that support emotional inferences, since the direction of influence could be the other way round or another factor could be influencing all of the skills, e.g. general intelligence. A longitudinal design or a study including training would be necessary to answer the question of causality.

Despite these limitations, our studies have still made a contribution to the research on situation models and emotional inferences, since they are the first studies examining children's emotional inference with an online measure. The results show that children from the age of 5 draw emotional inferences online. We can therefore assume that even preschoolers represent the protagonist's emotional state in their situation model, even if they are not explicitly asked to pay special attention to emotions. As mentioned at the beginning of this paper, being able to represent the protagonist's emotional state means a great learning opportunity about emotion from media. It also gives rise to a variety of new research questions. For example, do media have an effect on children's emotional knowledge? How important is this influence compared with other influences, such as family talk about emotions, peer group, school education, etc.? What does this suggest about the selection of TV shows and audiobooks suitable for different age groups? The results of our studies also showed that the ability to build emotional inferences depends on several personal factors,

such as age, emotional knowledge and media literacy. Although our studies constitute basic research, their results can be used as a basis for applied research in the educational sciences.

References

Baddeley, A., & Hitch, G. J. (1974). Working Memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). New York: Academic Press.

Barsalou, L. W. (1999). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes, 28*(1), 61–80.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645.

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior, 9*(2), 84–94.

Constantinidou, F., Danos, M. A., Nelson, D., & Baker, S. (2011). Effects of modality presentation on working memory in school-age children: Evidence for the pictorial superiority hypothesis. *Child Neuropsychology, 17*(2), 173–196.

Domaratius, S., & Ohler, P. (2006). *Ein Onlinetest zur Messung medialer Zeichenkompetenz bei Kindern*. 45th Congress of the Deutsche Gesellschaft für Psychologie, Nürnberg.

Dorr, A. (1998). Contexts for experience with emotion, with special attention to television. In M. Lewis & C. Saarni (Eds.), *The socialization of emotion* (pp. 55–85). New York, NY: Plenum Press.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues.* Oxford: Prentice-Hall.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Fecica, A. M., & O'Neill, D. K. (2010). A step at a time: Preliterate children's simulation of narrative movement during story comprehension. *Cognition, 116*(3), 368–381.

Feierabend, S., & Klingler, W. (2008). Was Kinder sehen. Eine Analyse der Fernsehnutzung Drei- bis 13-Jähriger 2007. *Media Perspektiven, 4*, 190–204.

Flavell, J. H., Green, F. L., & Flavell, E. R. (2000). Development of children's awareness of their own thoughts. *Journal of Cognition and Development, 1*(1), 97–112.

Fletcher, C. R. (1994). Levels of representation in memory for discourse. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 589–607). San Diego, CA: Academic Press.

Gernsbacher, M. A. (1995). Activating knowledge of fictional characters' emotional states. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 141–155). Hillsdale, NJ: Lawrence Erlbaum.

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. (1992). Do readers mentally represent characters' emotional states? *Cognition & Emotion, 6*(2), 89–111.

Gernsbacher, M. A., Hallada, B. M., & Robertson, R. R. (1998). How automatically do readers infer fictional characters' emotional states? *Scientific Studies of Reading, 2*(3), 271–300.

Gernsbacher, M. A., & Robertson, R. R. (1992). Knowledge activation versus sentence mapping when representing fictional characters' emotional states. *Language and Cognitive Processes, 7*(3-4), 353–371.

Gernsbacher, M. A., Robertson, R. R., Palladino, P., & Werner, N. K. (2004). Managing mental representations during narrative comprehension. *Discourse Processes, 37*(2), 145–164.

Gibbons, J., Anderson, D. R., Smith, R., Field, D. E., & Fischer, C. (1986). Young children's recall and reconstruction of audio and audiovisual narratives. *Child Development, 57*(4), 1014–1023.

Gillioz, C., Gygax, P., & Tapiero, I. (2012). Individual differences and emotional inferences during reading comprehension. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 66*(4), 239–250.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*(1), 1–55.

Gnepp, J., & Gould, M. E. (1985). The development of personalized inferences: Understanding other people's emotional reactions in light of their prior experiences. *Child Development, 56*(6), 1455–1464.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. In J. T. Spence, J. M. Darley, & D. J. Foss (Eds.), *Annual Review of Psychology* (Vol. 48, pp. 163–189). Palo Alto, CA: Annual Reviews Inc.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371–395.

Graesser, A. C., & Zwaan, R. A. (1995). Inference generation and the construction of situation models. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 117–139). Hillsdale, NJ: Lawrence Erlbaum.

Groeben, N. (2002). Dimensionen der Medienkompetenz: Deskriptive und normative Aspekte. In N. Groeben & B. Hurrelmann (Eds.), *Medienkompetenz. Voraussetzungen, Dimensionen, Funktionen* (pp. 160–197). Weinheim: Juventa.

Grüninger, C., & Lindemann, F. (2000). *Vorschulkinder und Medien.* Opladen: Leske & Budrich.

Gygax, P., Oakhill, J., & Garnham, A. (2003). The representation of characters' emotional responses: Do readers infer specific emotions? *Cognition & Emotion, 17*(3), 413–428.

Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. *Cognition & Emotion, 3*(4), 379–400.

Harter, S., & Buddin, B. J. (1987). Children's understanding of the simultaneity of two emotions: A five-stage developmental acquisition sequence. *Developmental Psychology, 23*(3), 388–399.

Hitch, G. J., Halliday, M. S., Dodd, A., & Littler, J. E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology, 7*, 347–362.

Janke, B. (2002). *Entwicklung des Emotionswissens bei Kindern.* Göttingen: Hogrefe.

Janke, B. (2008). Emotionswissen und Sozialkompetenz von Kindern im Alter von drei bis zehn Jahren. *Empirische Pädagogik, 22*(2), 127–144.

Johnston, W. A., & Hawley, K. J. (1994). Perceptual inhibition of expected inputs: The key that opens closed minds. *Psychonomic Bulletin & Review, 1*(1), 56–72.

Kail, R. (2000). Speed of information processing: Developmental change and links to intelligence. *Journal of School Psychology, 38*(1), 51–61.

Kelter, S., Kaup, B., & Claus, B. (2004). Representing a described sequence of events: A dynamic view of narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 451–464.

Kendeou, P., Bohn-Gettler, C., White, M. J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31*(3), 259–272.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York, NY: Cambridge University Press.

Koppenol-Gonzalez, G. V., Bouwmeester, S., & Vermunt, J. K. (2012). The development of verbal and visual working memory processes: A latent variable approach. *Journal of Experimental Child Psychology, 111*(3), 439–454.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology, 25*(6), 943–951.

Levine, L. J. (1995). Young children's understanding of the causes of anger and sadness. *Child Development, 66*(3), 697–709.

MacWhinney, B. (Ed.). (1999). *The emergence of language.* Mahwah, NJ: Lawrence Erlbaum Associates.

Markham, R., & Adams, K. (1992). The effect of type of task on children's identification of facial expressions. *Journal of Nonverbal Behavior, 16*(1), 21–39.

Marmolejo, R., & Heredia, A. T. (2006). Inferencias, modelos de situación y emociones en textos narratives: El caso de los niños de edad preescolar. *Revista Intercontinental de Psicología y Educación, 8*(2), 93–138.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition, 32*(8), 1389–1400.

Michalczyk, K., Malstädt, N., Worgt, M., Könen, T., & Hasselhorn, M. (2013). Age differences and measurement invariance of working memory in 5- to 12-year-old children. *European Journal of Psychological Assessment, 29*(3), 220–229.

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin, 135*(5), 749–773.

Molinari, C., Burin, D., Saux, G., Barreyro, J. P., Irrazabal, N., Bechis, M. S., …, Ramenzoni, V. (2009). Fictional characters' emotional state representation: What is its degree of specificity? *Psicothema, 21*(1), 9–14.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*(2), 319–334.

Munk, C., Diergarten, A. K., Nieding, G., Ohler, P., & Schneider, W. (2012a). Die kognitive Entwicklung des Verständnisses filmischer Montagetechniken bei Kindern als ein Kernbereich der Medienkompetenz. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 44*(2), 81-91.

Munk, C., Rey, G. D., Diergarten, A. K., Nieding, G., Schneider, W., & Ohler, P. (2012b). The cognitive processing of film cuts of 4- to 8-year-old children – an eye tracker experiment. *European Psychologist, 17*(4), 257–265.

Nieding, G. (2006). *Wie verstehen Kinder Texte? Die Entwicklung kognitiver Repräsentationen.* Lengerich: Pabst.

Nieding, G., & Ohler, P. (2004). Laborexperimentelle Methoden. In R. Mangold, P. Vorderer, & G. Bente (Eds.), *Lehrbuch der Medienpsychologie* (pp. 355–376). Göttingen: Hogrefe.

Nieding, G., & Ohler, P. (2008). Mediennutzung und Medienwirkung bei Kindern und Jugendlichen. In B. Batinic & M. Appel (Eds.), *Medienpsychologie* (pp. 379-402). Heidelberg: Springer.

Ohler, P. (1994). *Kognitive Filmpsychologie. Verarbeitung und mentale Repräsentation narrativer Filme.* Münster: MAkS.

Ohler, P., Nieding, G., & Töpper, J. (2002). Die Repräsentation von Protagonistenemotionen bei der Enkodierung narrativer Filme. In M. Baumann, A. Keinath, & J. F. Krems (Eds.), *Experimentelle Psychologie* (p. 70). Regensburg: Roderer.

O'Neill, D. K., & Shultis, R. M. (2007). The emergence of the ability to track a character's mental perspective in narrative. *Developmental Psychology, 43*(4), 1032-1037.

Palmer, S. (2000). Working memory: A developmental study of phonological recoding. *Memory, 8*(3), 179–194.

Pezdek, K., Simon, S., Stoeckert, J., & Kiely, J. (1987). Individual differences in television comprehension. *Memory & Cognition, 15*(5), 428–435.

Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C).* Hove: The Psychological Corporation.

Pons, F., Harris, P. L., & Rosnay, M. de. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology, 1*(2), 127–152.

Potter, W. J. (1998). *Media literacy* (1st ed.). Thousand Oaks, CA: Sage Publications.

Potter, W. J. (2011). *Media literacy* (5th ed.). Thousand Oaks, CA: Sage Publications.

Rapp, D. N., Gerrig, R. J., & Prentice, D. A. (2001). Readers' trait-based models of characters in narrative comprehension. *Journal of Memory and Language, 45*(4), 737–750.

Rey, G. D. (2012). *Methoden der Entwicklungspsychologie. Datenerhebung und Datenauswertung*. Norderstedt bei Hamburg: BoD.

Rideout, V. J., & Hamel, E. (2006). *The media family: Electronic media in the lives of infants, toddlers, preschoolers and their parents.* Menlo Park, CA: Kaiser Family Foundation.

Roseman, I. J., Antoniou, A. A., & Jose, P. E. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition & Emotion, 10*(3), 241–278.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes, 33*(2), 105–132.

Segers, E., Verhoeven, L., & Hulstijn-Hendrikse, N. (2008). Cognitive processes in children's multimedia text learning. *Applied Cognitive Psychology, 22*(3), 375–387.

Stein, N. L., & Levine, L. J. (1989). The causal organisation of emotional knowledge: A developmental study. *Cognition & Emotion, 3*(4), 343–378.

Stoneman, Z., & Brody, G. H. (1983). Immediate and long-term recognition and generalization of advertised products as a function of age and presentation mode. *Developmental Psychology, 19*, 56–61.

Subrahmanyam, K. (2008). Media symbol systems and cognitive processes. In S. L. Calvert & B. J. Wilson (Eds.), *Handbooks in communication and media. The handbook of children, media, and development* (pp. 166–187). Boston, MA: Wiley-Blackwell.

Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development, 79*(4), 866–881.

Therriault, D. J., & Rinck, M. (2007). Multidimensional situation models. In F. Schmalhofer & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 311–327). Mahwah, NJ: Lawrence Erlbaum Associates.

Tibus, M., Heier, A., & Schwan, S. (2013). Do films make you learn? Inference processes in expository film comprehension. *Journal of Educational Psychology, 105*(2), 329–340.

Tracy, J. L., Robins, R. W., & Lagattuta, K. H. (2005). Can children recognize pride? *Emotion, 5*(3), 251–257.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Unsöld, I. H. (2008). *Die Bildung von Inferenzen bei der kognitiven Verarbeitung medialer Texte.* Hamburg: Dr. Kovač.

Unsöld, I. H., & Nieding, G. (2009). Die Bildung prädiktiver Inferenzen von Kindern und Erwachsenen bei der kognitiven Verarbeitung audiovisueller und auditiver Texte. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 41*(2), 87–95.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York, NY: Academic Press.

Vygotsky, L. S. (1978). *Mind in society. Development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). A meta-analysis of theory of mind development: the truth about false belief. *Child Development, 72*, 655–684.

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development, 23*(2), 291–312.

Widen, S. C., & Russell, J. A. (2010a). Children's scripts for social emotions: Causes and consequences are more central than are facial expressions. *British Journal of Developmental Psychology, 28*(3), 565–581.

Widen, S. C., & Russell, J. A. (2010b). Differentiation in preschoolers' categories of emotion. *Emotion, 10*, 651–661.

Ziegler, F., Mitchell, P., & Currie, G. (2005). How does narrative cue children's perspective taking? *Developmental Psychology, 41*(1), 115–123.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*(5), 292–297.

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science: A Multidisciplinary Journal, 28*(4), 611–619.

Appendix A. Subscales of the Media Literacy Questionnaire with sample questions.

Subscale	**Example**
Knowledge about media and reality-fiction distinction	Does SpongeBob really exist?

Communication about media content	If you watch TV with somebody else, do you talk about what you see?
Enjoyment	Do you enjoy reading?
Reception	If you have to prepare a presentation for school, where do you look for information?
Selection and combination	Do you sometimes listen to music to make yourself feel better?
Media criticism	Is there sometimes stuff on TV that you don't like? Can you give me an example?
Productive participation	Have you ever taken pictures with a camera?

Appendix B. Subscales of the online measure for media sign literacy with sample questions.

Subscale	**Example**
Understanding of montage and continuity in films	"Which picture shows how the movie would continue?"
Differentiation between reality and fiction	"In which one of these films did the man and the woman really get married?"
Differentiation between filmic genres	Differentiation between fiction, news, and advertisements. "Which one of the three is not an advertisement?"
Understanding physical relations in comics	"Which one of these lions is running the fastest?"
Attributing voices of audiobooks	The voices differ in pitch, volume and speed. The

	participants are supposed to state which one of the shown figures (e.g. a fairy vs. a dwarf vs. a wizard) is talking.
Pictorial depictions of emotions	E.g. “smileys” and “frownys”
Detection of emotions in comic faces	“Which one of the persons is happy?”
Perspective taking	“Was this photo taken by the tiger, by the elephant or by the ant?”
Understanding of everyday symbols	E.g. colors. “Which one of these traffic-lights tells you to stop?”
Understanding of symbols used in maps	“Show me the mountain on this map”
Understanding of symbols used in computer programs	“Where would you click if you wanted to exit the game?”