Learning with Pictures and Accompanying Audio-Texts: Conflict Regulation, Narrative Processing, and Aesthetic Evaluation

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Stefan Manuel Knoos

aus Stuttgart

Tübingen

2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	07.07.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Stephan Schwan
2. Berichterstatter:	Prof. Dr. Katharina Scheiter

Table of Contents

Acknowledgements

I would like to thank my supervisors Dr. Manuela Glaser und Prof. Dr. Stephan Schwan for their advice, support and words of motivation. I learned a lot from you that was helpful for my academic research but also for my future career. I want to thank all colleagues in AG2 and the IWM for the warm and supporting working environment. Special thanks to Magdalena and to Ann-Kathrin for sharing the office, spending coffee breaks together, and sharing the experience of struggle and success as a PhD student. I would like to thank Melissa for being a great mentor when I started my PhD and also to many other PhD students at the IWM for sharing expertise on code, conferences, and formal stuff and spending entertaining breaks together. It was a great time working with you! I want to thank all research assistants for helping with data collection and creating the challenging audio-texts. I want to thank Dr. Schüler for reading my reports and providing helpful feedback. Thanks to all from the IWM Medientechnik for their help on audio-recordings, in setting up eye-tracking equipment, software and for the help in programming experimental environments. Lastly, I want to thank my family and friends. My family for their support in so many ways I cannot even count. Johanna a special thanks to you. Thank you Marsl and Anthea for proofreading. Many thanks to my friends in Vienna. There are many more persons I did not mention and who deserve a great thank you from my side: Thank you!

Summary

History paintings are a special form of pictures, as they can be used not only as historical documents, but also tell a story and are exhibited in galleries as artworks. However, history paintings often contain discrepancies between their depiction and reality that are frequently named in accompanying audio-texts. The aim of the present studies was to examine theories from three fields (picture as 1. historical document, 2. narrative, 3. artwork). These theories assume that naming of discrepancies between a picture and reality should affect the processing of the picture, the memory for content of the picture, and the evaluation of the picture. Study 1 showed that recipients fixated the content of the picture longer when it was named with discrepancies than without. This indicates that conflicts were recognized by the viewers. With regard to conflict regulation strategies, recipients did not process source information longer, nor did they report higher interest in further information when discrepancies were named than when discrepancies were not named. Thus, the results of Study 1 do not support neither sourcing nor information seeking as conflict regulation strategies of the recipients. In Study 2, recipients rated the trustworthiness of the picture document lower when discrepancies were named but not explained compared to when discrepancies were not named. As expected, explaining the discrepancies through the benevolent intentions of the artist compensated for the negative effect of naming discrepancies. This suggests that viewers used information about the source along with information about the content to evaluate the trustworthiness of the picture document. The recall of discrepant picture content was not better when discrepancies were named than without naming of discrepancies, but an exploratory analysis revealed that the recall of consistent pictorial content was worse when discrepancies were named. In Study 3, on the aesthetics of the picture, recipients did not rate their subjective understanding worse when discrepancies were named and not explained compared to when discrepancies were not named. However, aesthetic liking of the picture was significantly lower when discrepancies were named and not explained compared to not naming discrepancies at all. Moreover, as expected, this effect was compensated by an explanation of the discrepancies. There was

neither an effect of discrepancies nor explanation on surprise, interest, and confusion. The results show that fluency is important for aesthetic liking. Studies 1,2,3 did not show consistent results on whether naming discrepancies has a negative effect on the viewers' transportation. In Study 4, an audio text annotated two different picture documents about the same event. Discrepancies between the pictures were always named and differences in the artists' intentions were manipulated. Either artist's were described with differences in benevolence (one with propagandistic intention and one with a documentary intention) or without difference in benevolence (either both with a propagandistic intention or both with no intentions at all) The results of Study 4 did not show that recipients engaged in more sourcing and less information seeking when the source information was useful for conflict regulation than when the source information was not useful for conflict regulation. The present work extends theories in the respective fields and serves a better understanding of the design of audio texts in museums.

Zusammenfassung

Historienmalereien sind eine besondere Form von Bildern, da sie nicht nur als historisches Dokument genutzt werden können, sondern auch eine Geschichte erzählen und in Galerien als Kunstwerke ausgestellt werden. Häufig enthalten Historienmalereien jedoch Abweichungen zwischen Abbildung und Realität, die häufig in Audiotexten genannt werden. Ziel der vorliegenden Arbeit war es, Theorien aus den drei Feldern (Bild als 1. Historisches Dokument, 2. Narrativ, 3. Kunstwerk) zu untersuchen. Diese Theorien nehmen an, dass sich eine Benennung der Diskrepanz zwischen Bild und Realität auf die Verarbeitung des Bildes, das Gedächtnis für Bildinhalte und die Bewertung auswirken sollte. Studie 1 zeigte, dass Rezipienten Bildinhalte länger fixierten, wenn diese mit Diskrepanz benannt wurden, als wenn Diskrepanz nicht benannt wurde. Dies deutet darauf hin, dass Konflikte durch die Betrachter erkannt wurden. Hinsichtlich der Konfliktregulationsstrategien verarbeiteten die Rezipienten die Informationen über die Quelle weder länger, noch berichteten sie über ein höheres Interesse an weiteren Informationen, wenn Diskrepanzen genannt wurden, als wenn keine Diskrepanzen genannt wurden. Somit sprechen die Ergebnisse von Studie 1 weder für sourcing noch information seeking als von den Betrachtern verwendete Konfliktregulationsstrategien. In Studie 2 bewerteten die Rezipienten die Vertrauenswürdigkeit des Bilddokuments niedriger, wenn Diskrepanzen benannt, aber nicht erklärt wurden, im Vergleich zur Nichtbenennung von Diskrepanzen. Wie erwartet, kompensierte die Erklärung der Diskrepanz durch die wohlgemeinte Absicht der Quelle den negativen Effekt der Nennung von Diskrepanzen. Dies deutet darauf hin, dass die Betrachter die Informationen über die Quelle zusammen mit den Informationen über den Inhalt nutzten, um die Vertrauenswürdigkeit des Bilddokuments zu bewerten. Die Erinnerung an den diskrepanten Bildinhalt war bei Benennung von Diskrepanzen nicht besser als ohne, aber eine explorative Analyse ergab, dass die Erinnerung an den konsistenten Bildinhalt schlechter war, wenn Diskrepanzen benannt worden waren. In Studie 3, zur Ästhetik des Bildes, schätzten die Rezipienten ihr subjektives Verständnis nicht schlechter ein, wenn Diskrepanzen benannt und nicht erklärt wurden, im Vergleich zur Nichtbenennung von

Diskrepanzen. Allerdings war das ästhetische Gefallen des Bildes signifikant schlechter, wenn Diskrepanzen benannt und nicht erklärt wurden, im Vergleich zu nicht benannten Diskrepanzen. Dieser Effekt wurde zudem wie erwartet durch eine Erklärung der Diskrepanzen kompensiert. Weder die Benennung von Diskrepanzen noch deren Erklärung hatten einen Effekt auf Überraschung, Interesse oder Verwirrung. Die Ergebnisse zeigen, dass fluency für das ästhetische Gefallen wichtig ist. In Studie 1,2 und 3 zeigten sich keine einheitlichen Ergebnisse in Bezug darauf, ob sich eine Benennung von Diskrepanz negativ auf die transportation der Betrachter auswirkt. In Studie 4 kommentierte ein Audiotext zwei verschiedene Bilddokumente zum gleichen Ereignis. Dabei wurden immer Diskrepanzen zwischen den Bildern benannt und Unterschiede in der Intention der Künstler manipuliert. Entweder wurden beide Künstler mit Unterschieden im Wohlwollen ihrer Absichten beschrieben (einer mit einer propagandistischen Absicht der andere mit einer dokumentarischen Absicht) oder beide Künstler wurde ohne Unterschied beschrieben (Entweder beide mit einer propagandistischen Absicht oder beide ohne jegliche Absichten). Die Ergebnisse von Studie 4 sprechen weder dafür, dass die Rezipienten mehr sourcing betrieben, noch dass sie weniger information seeking als Konfliktregulation betrieben, wenn die Quelleninformation für die Konfliktregulierung nützlich war, als wenn die Quelleninformation für die Konfliktregulierung nicht nützlich war. Die vorliegende Arbeit erweitert Theorien in den jeweiligen Feldern und dient einem besseren Verständnis für die Gestaltung von Audiotexten in Museen.

1. Introduction

Information is not only documented, communicated, and acquired in the form of texts but often in the form of pictures. Pictures are used for these purposes throughout history, from cave paintings, over roman mosaics, medieval book illustrations, to modern paintings and photography. Pictures are useful historical documents as they can record acts of eye-witnessing and thereby can provide direct evidence about historic events. In addition, some events are only documented with pictures as for example hunting practices in ancient Egypt. Using picture documents augments our understanding of past events beyond merely using texts. Pictures are not only produced to convey information but also to persuade the viewers or give pleasure and thereby provide hints at past forms of knowledge, belief, and delight that texts wouldn't reveal. Pictures in contrast to texts allow to see the past more vividly. However, information from a picture cannot always be easily derived by merely looking at a picture's content. Although representational pictures, such as photography or representational paintings seem highly realistic, they frequently distort the reality (Burke, 2001). Propaganda films and photography of the totalitarian regimes of the 20th century clearly illustrate this point. Today, fake news remains a persistent problem and is provided in the form of text as well as in the form of pictures. These pictures are directly manipulated to convey false facts. But even pictures that were not intended to convey false facts must present the fact in a certain perspective. For example, national newspapers often depict their government's leaders in the center of the action of international events. To avoid being seized by perspectives and distortions, viewers need to consider context (van Boxtel & van Drie, 2012), and information about the artists (Burke, 2001). This is analogous to the argument that texts often represent one sided claims and readers must consider the content together with information about the author (Wineburg, 1991). Pictorial information is not only relevant in the field of history but can also represent debated scientific facts, such as diagrams plotting carbon dioxide and world temperature increases since the industrialization. Also, in the context of consumer products pictures are widespread for advertisement, product recommendations, or consumer protection. Although considering picture content together

with source information is relevant in nearly all disciplines, the present studies focus on realistic pictures of historic events relevant for history learning namely history paintings.

History paintings are a category of pictures that often contain distortions. They were highly valued during the nation-building era in Europe between the 18th and 20th century, before photography and film were invented. The artists aimed at depicting historic events realistically but were also influenced by their time, culture, and political interests. Therefore, the artists frequently and presumably intentionally included discrepancies between the content of the painting and the real historic event, as it can be assumed based on today's historians' opinions (Burke, 2001). Hence, when viewing multiple documents about the same historic event including history paintings, conflicting claims from the different artists and authors of these documents can frequently be noticed. Nowadays many history paintings are exhibited in museums of art and history. Museums often present these paintings together with explanations by a personal or audio guide, describing and interpreting the paintings' content, the historic event and today's historians' opinions on the distortions or inaccuracies of the paintings. In formal educations, history paintings, historical photography, or historical film documents are explained and discussed by teachers. Also, in other disciplines and learning situations, picture documents are often accompanied by verbal explanations such as in talks, presentations, or news reports. However, independently of all these learning scenarios, it is often difficult to understand picture documents on their own and therefore, accompanying verbal explanations need to be provided in written or audio format.

1.1 The Role of Audio-Texts for the Understanding of Pictures

Audio texts play an important role for the understanding of pictures, since acquiring information from complex pictures is difficult due to differences between the processing of text and picture information. First, written text information is processed linearly, in Western cultures from left to right and top to bottom, but pictorial information is not processed in the same sequence by all viewers (Larkin & Simon, 1987). When processing pictures, the focus is on salient and central elements (Tatler, 2007) and important non-salient or peripheral

pictorial information is easily missed by the viewer. Second, pictorial information is often ambiguous (Pozzer-Ardenghi & Roth, 2005). Audio-text may help to direct the viewer's attention to important parts of the pictures and disambiguate the information. Based on the consideration of differences between text and pictures and theories of multimedia processing (Mayer, 2009), it is assumed that an extended multimedia effect applies for the processing and understanding of pictures (Glaser & Schwan, 2015). The multimedia effect (Mayer, 2009) describes that information from text is learned better when accompanied by a picture. In contrast, for the extended multimedia effect, a picture is considered as the central information resource instead of the text. The extended multimedia effect states that pictorial information is better learned when accompanied by textual information than without additional textual information. This might be due to dual coding (Paivio, 1990) but also due to the text guiding the viewer's attention on relevant pictorial information (Glaser & Schwan, 2015). Research showed that naming the pictorial elements of architecture in a simultaneously presented audio-text lead to longer fixation times on these elements and better memory for the content compared to unnamed pictorial elements (Glaser & Schwan, 2015). This positive effect of naming pictorial elements in an accompanying audio text on the processing and the memory of pictorial elements was also shown for salient and non-salient pictorial elements of historical paintings (Glaser et al., 2020). Thereby the pictorial elements were named and interpreted with regard to the historic event.

As already noted above, pictures do not always truly represent the historic event. Other documents can provide additional information about the event differing from the representation made by the picture. For example, the painting about Washington's crossing of the Delaware by the artist Emanuel Leutze depicts the scene in a bright daylight (1851, Metropolitan Museum of Art, New York City). In contrast to the painting, an audio text could provide information that based on today's historians' opinions it was nighttime during the crossing of the Delaware. For such a situation, multimedia theory provides a limited theoretical approach, since it focuses on situations where text and picture provide coherent and non-conflicting information. In addition, different authors of texts and pictures are not

relevant in multimedia research. Therefore, additional theories need to be considered for a situation where text and picture documents provide conflicting perspectives from different authors.

According to the theories presented in the following chapters, conflicting information affects three aspects relevant for viewers dealing with pictures. First, the picture as a document. Second, the picture as a narrative telling a story, and third, the aesthetic experience of the picture as an artwork. History paintings are a category of pictures for which all three aspects are highly relevant, while other pictures may only be documents but do not tell a story or are not considered and labeled as artworks. In the following section I describe theories and empirical evidence about the effects of conflicting information on these three aspects of pictures: the pictures as a document, a narrative, and an artwork. Most of the theories in the context of multiple documents and narrative processing focus on texts but explicitly or implicitly apply to content in other media formats than texts. The theories in the context of aesthetics apply to all artworks but are mostly investigated with paintings.

1.2 Conflict Detection and Regulation with Multiple Conflicting Documents

How readers handle conflicting information is a central theme in multiple document research. When reading more than one text about a complex topic it is often the case that different authors propose different arguments and opinions about the topic. The document model (Britt et al., 1999; Perfetti et al., 1999) extends theories of single text comprehension to fit this situation. According to the document model and adopted from a single text comprehension model (Kintsch, 1988) readers first construct three layers of representation: a surface code, a text-base, and a situation model. The surface code is the exact wording of sentences, which is stored only for the most recent sentences. The text-base incorporates the meaning of sentences including a network of propositions but does not incorporate the exact wording. The situation model contains what the whole text is about together with inferences from background knowledge. In addition, for the mental representation of multiple documents, two more layers are needed: First, an integrated mental model has to be built

which is a global representation about the topic constructed by integrating the situational models of the different documents. The second additional layer that has to be generated is the intertext model. It includes document nodes and links between the nodes. A document node represents information about the author, the rhetorical goals and the form. Readers also construct links between different document nodes (e.g. author A and B disagree) and links from a document node to the content (e.g. author A said …). Readers do not always construct such links between content and source since this is a cognitively demanding task. However, conflicting information seems to enhance the readers engagement in sourcing (Braasch & Bråten, 2017), since representing source information in document nodes allows the reader to construct a coherent mental model when the authors provide conflicting information. The process of attending author information, evaluating, and representing it is referred to as sourcing.

Sourcing is one conflict regulation strategy also described by the content-source-integration model (CSI; Stadtler & Bromme, 2014).The CSI model describes more detailed the readers' conflict regulation strategies when conflicting information from two or more documents is encountered. The model assumes three stages of conflict regulation. In the first stage, a conflict needs to be detected by the reader. In the second stage, the conflict needs to be regulated by the reader in order to restore coherence. To restore coherence, readers can apply different strategies. First, they can ignore the conflict. Ignoring is defined by a processing behavior that indicates that readers have detected a conflict but do not remember it afterwards. Second, readers can restore coherence also by reconciling the conflicting claims. This means they wrongly remember the conflicting claims as if they were not conflicting but coherent statements. Third, readers can draw inferences to explain the conflicts. Lastly, they can accept the conflict due to different sources. In the third stage, of conflict resolution, readers can try to resolve the conflict between claims of two different sources. If readers have high prior knowledge, they might be able to answer the question what is true but if this does not apply to the readers, they can resolve the conflict only by asking whom to believe. To answer the latter question, readers need to judge the

trustworthiness of a document and its source (cf. Saux et al., 2018). Hence sourcing is one important strategy to regulate and reconcile conflicting information from multiple documents.

Besides the resolution mechanisms described by the CSI-model, readers can also postpone the conflict resolution if the given information is not sufficient to resolve the conflict and wait or seek for future information to resolve the conflict (Graesser et al., 1994). Thereby the recipient recognizes a knowledge gap and becomes more interested in receiving further information to finally resolve the conflict and the gap in knowledge (Loewenstein, 1994; Murayama et al., 2019). Empirical evidence supports this assumption by showing that readers ask more information seeking questions after encountering conflicting information (Graesser & Olde, 2003; Otero & Graesser, 2001). The following chapters present empirical evidence supporting assumptions of the document model and the three stages of conflict processing proposed by the CSI model, namely conflict detection, regulation, and resolution. First, effects on processing and memory are described resulting from conflict detection (chapter 1.2.1). Then, empirical evidence on sourcing is described as a conflict regulation strategy in order to restore coherence and find a resolution for the conflict (chapter 1.2.2). Lastly, findings of studies using multiple text documents and mixed documents (texts and videos) are described. Based on this, potential challenges for generalization from findings of multiple text document research to multiple documents using other media formats than written text are outlined (chapter 1.2.3).

1.2.1 Empirical Evidence on Processing and Memory Due to Conflict Detection

Single text studies frequently show that readers detect conflicts between two parts of a written text (Albrecht & Myers, 1995; Albrecht & O'Brien, 1993; O'Brien et al., 1998; O'Brien & Myers, 1985). The detection of conflicts is indicated by longer reading times of a sentence that conflicts with previously read sentences compared to if the same sentence is not in conflict with prior text. These single text studies often use narrative texts where the main character is described as for example vegetarian and later, this character does something that is either conflicting (e.g. ordering meat) or not (e.g. ordering salad) with the

previously described characteristic. The presence of conflicts in texts does not only increase the readers processing times but can also improve memory for the conflicting content at least if readers do not ignore or distort parts of the conflicting information (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995).

Only a few experiments investigated conflicts with other material than texts. One study (Experiment 1 by Schüler, 2017) combined pictures with written text. The text described a pictorial element of a map, such as its color, and was either consistent or in conflict with the depiction of the element in the picture. Conflicting compared to consistent information resulted in longer fixation times on the text and the picture, but worse memory for the pictorial content. In the second experiment by Schüler (2017), the mechanism of a toilet flush was used as learning content. The longer fixation times due to conflicting information was replicated, but only for the text content, not for the picture content. For the learning outcomes, no differences were observed between the conditions. Interestingly, about 80% of the participants did not remember the conflicts between the text and the picture afterwards. This combination of detection without memory for the conflict indicates that they might have ignored the conflict or distorted parts of the conflict in order to restore coherence. The results of a third study (Schüler, 2019) revealed that fixation counts were higher on the text and on the picture when these were in conflict with previously presented textual and pictorial information compared to when they were consistent. The participants in the two conditions did not differ regarding memory for the content. In this study two-thirds of the participants remembered the conflict, which might be explained by the information being presented twice to the participants.

In sum, these results about conflict detection using text-picture material are in line with empirical findings of single text studies (e.g., Albrecht & O'Brien, 1993). The findings of the text-picture studies strongly support the conflict detection stage of the CSI model (Stadtler & Bromme, 2014) and the processes of co-activation and information integration into a coherent mental model (Kintsch, 1988). In addition, the results indicate that viewers or

readers encountering conflicting information might engage in top-down processes to resolve the conflicts (Graesser et al., 1994) and might not always remember the conflicting information afterwards although they detected the conflict (Baker, 1979). Differences between single text studies and text-picture studies emerged with regard to memory. Single text studies (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995) showed better memory for conflicting compared to consistent text. In the text-picture studies content memory was worse (Experiment 1 of Schüler, 2017) or did not differ between the conditions with and without conflict (Experiment 2 of Schüler, 2017, 2019). One reason for this could be that the single text studies measured memory for the conflicting content while the text-picture studies measured memory not for conflicting content but for the content that was presented as consistent in all conditions.

1.2.2 Empirical Evidence on Sourcing as a Conflict Regulation Strategy

In the aforementioned studies information about the sources was not available. When source information is available, single text studies frequently show that readers engage in more sourcing when different sources provide conflicting claims than when they provide coherent claims (Braasch et al., 2012; Rouet et al., 2016; Saux et al., 2017). This was also shown for conflicting claims of sources from multiple text documents (Kammerer et al., 2016; Strømsø et al., 2013). Sourcing describes the process of attending source information, evaluating, and representing it. Attention on source information is thereby often measured by eye-tracking and results revealed longer processing times of source information when the documents present conflicting claims than when they do not present conflicting claims (Braasch et al., 2012; Braasch & Bråten, 2017; Kammerer et al., 2016). Readers encountering conflicts often have a better memory for links between authors and contents after reading than readers of texts without conflicting information (Braasch et al., 2012; Saux et al., 2017). During sourcing the readers evaluate the source information in order to judge reliability or trustworthiness of the source and hence the content of the document (Saux et al., 2018). This evaluation of sources can help the reader to explain the conflicts between the

claims by different authors (Stadtler & Bromme, 2014). One study showed that conflicting information increases the memory for source information that helps to explain the conflict, such as expert knowledge, but not for source information that is irrelevant for an explanation of the conflict, such as the appearance of an author (Saux et al., 2018). Two aspects of the source seem to be especially important for the evaluation of trustworthiness, namely expertise and benevolence (Thomm & Bromme, 2016). Benevolence is the readers appraisal of whether the author has good intentions for the audience and tries to provide the best possible information or whether the author is following selfish goals, such as commercial interests (Mascaro & Sperber, 2009; Thomm & Bromme, 2016). When conflicts were present in both conditions, differences in authors' benevolence lead to longer processing of the source information but not a better memory for source information compared to when the authors did not differ with regard to their benevolence (Gottschling et al., 2019). In another study participants better remembered the authors' names when the authors differed in expertise or benevolence compared to when they did not differ (Thomm & Bromme, 2016). In addition, claims by benevolent authors are rated more useful than claims by commercially biased authors (Kammerer et al., 2016) and the trustworthiness of authors is positively linked to the number of citations in a subsequent essay about the topic presented in the texts (List et al., 2017). As already mentioned most studies used multiple text documents. In the following I will describe several studies using multiple text, video, and mixed mode documents.

1.2.3 Empirical Evidence on Conflict Regulation with Multiple Video Documents and Documents in Different Media Formats

Several studies directly compared multiple document situations with either only texts, only videos, or combinations of text and video documents (List, 2018; List & Ballenger, 2019; Salmerón et al., 2020). Mixed results emerged for comprehension strategies used by the learners across different types of material. One study found that learners mostly used the same strategies to comprehend text and video documents and had a similar good

comprehension about the topic (List, 2018). Another study (List & Ballenger, 2019) showed that learners using multiple text documents reported more frequent use of comprehension strategies, such as cross-textual elaboration than learners using video documents. In addition, comprehension was better when the text document was presented prior to the video document compared to the opposite order. However, these studies did not consider conflicts between the documents and only general comprehension strategies but not specific conflict regulation strategies were examined.

In the study by Salmerón et al. (2020) conflicting documents were used informing about the advantages and disadvantages of drinking tapped or bottled water. Again, learners either studied the material in the form of two texts, two videos, or a combination of text and video. The results revealed that learners defended the claims made in the video more often than claims made in texts when they received mixed documents. Integration was better with two text documents than with two video documents, indicated by a higher number of inferences. Overall source memory was not affected by the different presentation modes, but learners better remembered the profession of one source, the doctor, if presented in the video than in texts. This might be explained by the doctor's profession being not only stated in the text but visibly cued by wearing a white coat in the video. This study not only shows that conflicts influence source memory but also how sources are presented by the possibilities of different media types could further affect the memory for the sources.

A recent study investigated how the way of presenting source information affects the recipients source memory and use to explain the conflict with combinations of texts and videos (Merkt & Huff, 2020). Experiment 1 of the study manipulated the spatial contiguity of the source and the documents by varying the position of the source information. All source information was either presented together before all documents or together after all documents or each source information was presented separately and directly before the respective document or separately and directly after the respective document. In an open-ended task asking recipients about potential reasons for commonalities and differences

between the documents, participants more often explained the conflict due to different sources when the source information was presented separately before the documents compared to separately after the documents. However, when the source information was provided all together, the participants who received the source information subsequent to the documents explained the conflict more often due to different sources compared to those participants who received the source information prior. The position of source information did not affect participants' construction of source content links, measured with a verification task. In Experiment 2 of the study by Merkt and Huff (2020) the source information was again presented prior versus following the presentation of documents, but participants were additionally either informed about the conflicts prior or subsequently to the documents. This did not affect the pattern of results observed in Experiment 1, indicating that these results cannot be explained by a lack of conflict awareness when the information about the author is presented before the conflict, as it would be assumed based on the D-ISC model (Braasch et al., 2012).

In sum, these studies hint at potential challenges and limitations for generalization of findings from multiple text documents to multiple documents of other media formats. For example, the format might determine the way source information is presented. In videos the source might be visible in the video and not only described by text, which could lead to better source memory (Salmerón et al., 2020) due to dual coding. Empirical findings and considerations in the context of multimedia research (Mayer, 2009) may help to describe potential differences with regard to conflict detection and regulation when using media formats different from written text documents. The non-linear processing of pictures (Larkin & Simon, 1987) might increase the chance that conflicting pictorial elements are overlooked. Higher transiency of audio compared to written text (Wong et al., 2012) could also lower the detection of conflicts as well as memory for source information. In contrast, concurrent presentation of text and picture documents leads to higher temporal contiguity (Ginns, 2006) and possibly fosters conflict detection. Therefore, the present studies aimed to investigate

whether effects known from multiple document literature using mainly written texts generalizes to text and picture documents.

1.3 Conflicting information and Narrative Processing

History paintings are inherently narrative and tell a story about the depicted event. According to the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008), readers or viewers construct a mental representation of a presented narrative. The primary mental model for comprehension of a story is the situational model. The situational model includes events and characters and is constantly updated during the processing of the story. According to the event-indexing model events are connected along five dimensions of the story: time, space, protagonist, causality and intentionality (Zwaan et al., 1995). For constructing a situational model, recipients draw on pre-held schemas about the story world and characters based on genre and stereotypes. When recipients can easily construct the mental model, they experience transportation, which is a flow-like state of being absorbed by the story. Different from just being absorbed, recipients locate themselves in the story by performing a deictic shift (Segal, 1995). Thereby the story world becomes the center of the recipients' experience.

Transportation consists of a cognitive, an emotional, and an imaginative component (Green & Brock, 2002). When recipients are transported, they focus their cognitive processing on the events described in the narrative, they identify with the characters in a story and experience similar emotions as them. Transported recipients also create vivid mental images of the places and characters of the story. Transported recipients lose awareness of what is going on in the real world around them up to the level that they have the feeling of being lost in a story (Nell, 1988). Narratives have a persuasive impact on the transported recipients and recipients might more readily believe claims presented within a narrative. They also may become less critical of information presented in the narrative and do not counterargue. Narrative processing differs from an elaborative processing. The former

is characterized by experiencing facts vividly and emotionally as if oneself is acting in the story world whereas the latter is characterized by logical analysis of the facts. Transportation is also enjoyed by the recipients (Green & Brock, 2000).

According to the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008), recipients of a narrative do not assess its realism unless they encounter clear violations of narrative realism or external realism. Violations of narrative realism refer to discrepancies between the facts presented within the story. Violations of external realism refer to discrepancies between the facts presented in the story and facts of the real world. Violations of narrative and external realism hamper narrative processing and the flow of constructing the mental model and thereby reduce the recipient's experienced transportation. Therefore, based on the model it can be predicted that discrepancies between the content of a painting and the real-world are violations of external realism and would lead to critical evaluation of the narrative presented in the painting and lower the experienced transportation and enjoyment.

1.3.1 Empirical Evidence on the Influence of Conflict Detection on Transportation

In one study participants read a newspaper article about an 8-year-old drug addict. After reading the article the participants were either provided with the information that the author of the article intentionally invented the story to become famous and presented the story as factual although it was not, or they were provided with the information that the author unintentionally provided invented facts, or they were not provided with any information before the dependent variables were measured. Results revealed that the newspaper article changed the readers' attitudes independently of whether the story was stated to be intentionally untrue, unintentionally untrue or presented without information about the author. In the notes the authors further state that transportation did not differ between the conditions and therefore the naming of seemingly realistic newspaper reports as untrue did not lower their transportation (Green & Donahue, 2011). The authors argue that this might have been because the wrong information was not identifiable by the readers who possessed only low

prior knowledge about the real world of drug consumers, and drug politics. When clearly naming which facts are wrong this might have been different. In addition, presenting recipients with alternative facts to the untrue facts could challenge the construction of a mental model of the story and lead to a lowering of transportation.

1.4 Conflicting Information and Aesthetic Evaluation

How additional information such as titles or text labels affects the evaluation of artworks is a prevalent question in the field of aesthetics with substantial empirical findings. Three theoretical frameworks are relevant for the research in this field. First, the fluency theory (Reber et al., 2004) assumes that the liking of artworks is linked to the viewer's ease of processing, meaning making, and understanding of an artwork. If additional information makes the viewer's processing easier, the artwork should be liked more by the viewer. In contrast, if additional information makes processing and understanding for the viewer more difficult the artwork should be liked less. Fluency evolves either during early processing stages on the perceptual level or later during higher order processing stages related to meaning making and understanding, such as the cognitive mastery of an artwork (Belke et al., 2010).

The second framework is the psycho-historical framework for the science of art appreciation (Bullot & Reber, 2013). The framework describes three modes of art appreciation. The first mode is basic exposure. Thereby, viewers perceptually explore the artwork, implicitly classify it according to its style and experience aesthetic emotions. In this mode the outcomes are assumed to be similar to those proposed by fluency theory and can be subsumed under the simplified hypothesis that a better and easier understanding of an artwork should be positively linked to its aesthetic liking. In line with this, a lower or difficult understanding results in the opposite. The psycho-historical framework proposes that disfluency should be liked less during the basic exposure, but the framework extends fluency theory with the assumption that disfluency in the context of art can trigger the second mode, the design stance. During the design stance viewers ask for more information about context

of the artwork such as the intentions of the artist and the history of the artwork to get a better understanding of the artwork. This second mode is more likely adopted by experts of art than laypersons. Also, experts more likely reach the third mode named artistic understanding. During the process of artistic understanding, viewers consider art historical information in order to understand the artwork more deeply. Thereby, viewers can start to like initial disfluency of an artwork if they can appreciate the disfluency as an artistic element. For example, artists can intentionally include disfluency with the purpose to provoke the recipients' analytical thinking. Since empirical research is often conducted with laypersons most studies investigate the simplified hypothesis of the framework that a better understanding leads to better liking (e.g. Swami, 2013). In sum, both fluency theory and the psycho-historical framework for the science of art appreciation claim that additional information, such as titles or verbal explanations of the style and the art historical context, can enhance but also lower the liking of artworks, depending on whether the additional information contributes to a fluent processing and better understanding or to a less fluent processing and lower understanding.

The third framework is the Vienna integrated model of art perception (VIMAP; Pelowski et al., 2017). This model tries to incorporate the two previous mentioned frameworks together with the model of aesthetic appreciation and aesthetic judgments (Leder et al., 2004) in a more complex model about art processing and aesthetic evaluation. The model proposes seven stages. The first stage is the pre-classification. During pre-classification viewers are affected by factors of the context of viewing the artwork (museum, laboratory, social or individual setting) and personal factors (mood, personality, expectations). In this stage the viewers do not even process the artwork itself. The second stage is the perceptual analysis of the low-level features of an artwork. Low level-features are for example complexity, contrasts, and color. During the third stage, the implicit memory integration, the viewers try to combine elements of the artwork to meaningful patterns. This is influenced by factors such as familiarity and prototypicality. In the fourth stage, the explicit classification, viewers identify the content in accordance with the artwork's context, style, and

the knowledge learned about the artist. These first four stages are mainly dominated by bottom-up processing.

The fifth stage is the cognitive mastery. Cognitive mastery is characterized by top-down processes that consider and combine the information collected during the previous four stages in order to form a coherent understanding of the artwork together with an appropriate evaluation and physical response. Two processing checks influence the outcome during the cognitive mastery stage: the schema congruency check and the self-relevancy check. For schema congruency, viewers check their schemas within their own knowledge, expectations, understanding, and possibilities for learning. Viewers also check whether they think the processing during the former stages of basic perceptual processing, object identification, explicit classification, and integrating was successful. Viewers rate the match for each of these elements on congruency. A good overall match results in a subjective feeling of fluency and an efficient processing and understanding. The second check of the model is the self-relevancy check. Thereby, viewers take into account the personal importance of the artwork or the personal importance of understanding the artwork for their self-image. Thereby, high importance of artworks and the outcome of the processing of artworks is mainly relevant for experts of art but not for laypersons, since, in contrast to experts, the outcome of the layperson's art processing does usually not threaten their self-image. Hence, when considering only the laypersons' art processing, the self-relevancy check can be ignored. In experiments with laypersons only the following two different outcomes of the model are possible. These two outcomes depend on the congruency-check.

High schema congruency together with low self-relevance results in a default or facile reaction. This is the most common outcome of naive viewers who glance at the artwork not finding something new or questioning. This results in a subjectively good enough classification, an easy processing and seemingly good understanding of the artwork. The viewers will experience little emotions and a default feeling of pleasure.

In contrast, low schema congruence together with low self-relevance results in a reaction of novelty and small insight together with some aesthetic emotions. Viewers should generally experience more surprise when incongruence is encountered than when no incongruence is encountered. The emotional reaction regarding interest, confusion, understanding and liking depends on whether or not the viewers feel the need to solve the incongruence and whether they are able to do so. Viewers can render the incongruence as irrelevant and not experience the need for a resolution. Viewers can also accept the incongruence as a mystery and tolerate the ambiguity and not seek a resolution. If the viewers feel the need for a resolution but appraise the chance of finding a resolution to be low, the viewers will experience confusion together with lower interest, and tendentially a lower subjective understanding and liking of the artwork. If the viewers feel the need for a resolution and appraise a good chance to do so, they can overcome incongruency. If the incongruence is resolved, viewers might again experience pleasure and interest. Viewers can resolve the incongruence either by continuing their processing to find more information that contributes to a higher match, or they modify their schema by generalizing definitions, classes, or expectations to include the novel elements. Lastly, viewers can wait or search for further information that explains the incongruence.

In sum, fluency theory proposes a link between fluency of processing or understanding and the aesthetic liking. The psycho-historical framework for the science of art appreciation extends this assumption by proposing that it is possible that disfluency of artworks is liked but this might be restricted to experts of art. The VIMAP proposes more detailed the emotional reactions and aesthetic evaluation of viewers including laypersons when schema-congruence is low.

1.4.1 Empirical Evidence in the Context of Aesthetic Evaluation

Research in the field of aesthetics mostly support the models by showing positive effects of additional information, such as titles or context information on the viewer's subjective understanding of paintings (Bubić et al., 2017; Leder et al., 2006; Russell, 2003;

Swami, 2013) and their aesthetic appreciation (Belke et al., 2010; Bubić et al., 2017; Gerger & Leder, 2015; Millis, 2001; Russell, 2003; Swami, 2013). In addition, the results of two of these studies support the assumptions of the theoretical models described above that additional information accompanying artworks can both positively as well as negatively affect the viewers aesthetic evaluation (Belke et al., 2010; Gerger & Leder, 2015). In these two studies the semantic match of a title with the content of the painting was manipulated. One group received titles matching with the depicted content, another group received non-matching titles and the control condition did not receive any titles. In both studies liking was highest when the title did match with the content and lowest when it did not match semantically with the content of the painting. To explain this effect, the authors refer to fluency theory and propose that liking was lower because of a lower fluency due to the inconsistencies between the title and the painting's content. Hence, these studies support that additional information can reduce the liking of paintings if they provide the viewer with inconsistencies.

Although there seems to be a strong support for effects of additional information on subjective understanding and liking of artworks, a recent review (Chmiel & Schubert, 2019) points out a substantial number of studies that did not observe these effects. The review points out several boundary conditions. One boundary condition seems to be iconicity of the painting. Some studies support positive effects of additional information on subjective understanding and liking for highly abstract but not for representational art (Leder et al., 2006; Moore & West, 2012; Swami, 2013). One explanation for this is the already high subjective understanding and liking of laypersons viewing representational paintings without any additional information. In contrast, the laypersons' understanding and liking of abstract art is low without additional information and easily increased by additional information. Hence, a ceiling effect may be the reason for the absence of effects of additional information on the understanding and liking of representational art.

Another difficulty is that the type of information influences the effect of additional information on subjective understanding and liking. One study compared different types of titles (Leder et al., 2006) and revealed that both descriptive and elaborative titles improved the subjective understanding of paintings but elaborative titles had the greatest effect compared to a control group without titles. However, neither of these types of titles increased liking of the paintings. Another study compared titles, broad genre information, and content specific information to a control group without additional information (Experiment 1 of Swami, 2013). The results of this experiment revealed that all three types of information improved the subjective understanding of abstract paintings compared to the control condition, but content specific information had the highest effect and was the only type of information that improved liking. Together with the findings that additional information accompanying artworks can also have negative effects on the viewers aesthetic evaluation (Belke et al., 2010; Gerger & Leder, 2015) the described studies provide a good support for the theories described above. However, this is only the case for abstract and not representational art and in addition the type of information should be considered. Content specific information may have greater effects on subjective understanding and liking compared to using only titles.

With regard to effects of additional information on aesthetic emotions other than liking, empirical evidence is scarce. Former studies mostly reported no effects of additional information in the form of titles on the emotional experience in general (Bubić et al., 2017) or interest (Gerger & Leder, 2015; Leder et al., 2006). To the best of my knowledge, no studies have investigated the effects that additional information can have on surprise or confusion.

In sum, studies revealed positive effects of additional information on subjective understanding and liking mostly for abstract but not representational art. Studies investigating negative effects of additional information on aesthetic liking only used titles but no longer verbal explanations. Similar to mismatching titles, but more common in the real world are discrepancies between the content of a painting and the reality it represents. Historical paintings frequently contain discrepancies between the depiction of a historic event

made by the artist and the real event as it can be assumed based on today's historians' opinions (Burke, 2001). These inaccuracies can be mentioned by audio-text commenting on the painting. Mentioning inaccuracies is important for interpreting and understanding the painting and is often done verbal explanation of personal or audio guides provided by museums of art and history. However, based on the theory it can be expected that naming inaccuracies affects subjective understanding and liking of the paintings negatively, at least when the inaccuracies remain unexplained.

1.5 The Role of Conflicts in Theories from the Fields of Multiple Documents, Narrative Processing, and Aesthetic Evaluation

The three models, CSI (Stadtler & Bromme, 2014), VIMAP (Pelowski et al., 2017) and the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008) contain similarities and differences with regard to the concept of conflicts and potential conflict regulation. The CSI focuses on conflicts between two or more different claims about a situation that cannot be true at the same time. The VIMAP describes incongruence as a match between the information processed and the schemas held by the viewer. The VIMAP's concept of incongruence is the broadest term and can arise from many different situations. It is reasonable that conflicts lead to a more difficult understanding and a lower congruency match (Pelowski et al., 2017). Therefore, conflicts in the sense of the CSI would also lead to incongruence in the sense of the VIMAP. Lastly, inconsistencies in the sense of the model of narrative comprehension and engagement are violations of either external realism or narrative realism. External realism is a match between the story and the actual world. Narrative realism refers to a plausibility and coherence within the story. Again, conflicting claims can also violate a narrative's external realism if clear inconsistencies between the facts presented in a story and the real world are debated. Regarding conflict regulation, the model of narrative comprehension and engagement does not propose any specific conflict regulation strategies but briefly claims that the violation of external and narrative realism only reduces transportation as long as these violations remain unexplained (Busselle & Bilandzic,

2008). While the VIMAP assumes that inconsistencies can lower aesthetic liking the model of narrative comprehension and engagement similarly assumes that enjoyment is lower when the experienced transportation is decreased.

The VIMAP and the CSI model share more communalities regarding conflict regulation, since both propose that recipients apply strategies to restore coherence. These strategies are highly similar. According to the VIMAP, viewers can render the incongruence as irrelevant and not experience the need for a resolution which is analogous to recipients ignoring the conflict as proposed by the CSI. The VIMAP assumes that viewers can adjust their schemas to fit incongruent information of an artwork, while the CSI assumes that readers can adjust one part of conflicting information to fit the other side of the conflict in order to restore coherence. The VIMAP describes that viewers can resolve the incongruence by continuing their processing to find more information that contributes to a higher match. In the case of the CSI this would specifically focus on the processing of source information to solve the questions what is true or whom to believe. Or as stated by other authors, the viewers may use the strategy to seek further information to resolve the conflict (Graesser et al., 1994). A major difference resulting from this strategy is the outcome of interest. In the context of learning with multiple documents I identified information seeking as a relevant strategy applied by recipients after conflicts were encountered (Graesser et al., 1994; Graesser & McMahen, 1993). The recipients perceive a knowledge gap and try to fill this gap (Murayama et al., 2019). Therefore, they become interested in receiving specific further information to resolve the conflict. In contrast, in the context of aesthetics interest is perceived as a state that emerges when incoming information is novel but easy to process and comprehend. Interest is thereby oppositional to confusion that results from novel information that is not easy to process and comprehend (Silvia, 2009). Considering both theories it is possible that conflicts lead to less general interest of the viewer with regard to the artwork but to higher interest in specific information to resolve the conflict and close the knowledge gap. A second major difference in strategies described by the VIMAP might be that for artworks there can also be a certain pleasure in mystery and viewers might tolerate

the ambiguity and not seek any resolution. In sum, although differences between the models emerge, they share that conflicting information is highly important for outcomes in the respective fields that should be investigated by empirical research.

1.6 Present research

In the present studies I investigated combinations of picture documents together with audio-texts. The audio-text thereby supported information acquisition by explaining the pictures and furthermore directly pointing out discrepancies between the pictures content and the real historic event based on today's historians' opinions. In my last study the audio-text pointed out discrepant claims between two picture documents depicting the same event but did not refer to the "real historic event". In both cases viewers are confronted with conflicting claims from two documents. In my investigations I considered relevant outcomes of conflicting information in three fields, namely multiple documents, narrative processing and aesthetics.

Research in the context of multiple documents mostly considered text documents. Some studies considered videos (List & Ballenger, 2019; Merkt & Huff, 2020; Salmerón et al., 2020) but to the best of my knowledge no study investigated static pictures together with audio-text and it is unknown whether effects known from text studies generalize to such different material using audio-text and picture. Differences between multiple documents of written text and documents of picture and audio-text emerge from the media format. For example, picture and audio text documents can be presented simultaneously instead of sequentially. This could lead to higher temporal contiguity (Ginns, 2006) and an easier detection of conflicts. In contrast, audio texts inherit a higher transiency of information than written texts (Wong et al., 2012). Based on these considerations, my first aim was to investigate whether effects known from the processing of conflicting texts generalizes to pictures together with audio-text.

My second aim was to investigate effects of conflicts on narrative processing. The model of narrative comprehension and engagement proposes that violations of external

realism would hamper the construction of the mental model and thereby lower transportation (Busselle & Bilandzic, 2008). The assumptions of the model of narrative comprehension and engagement are proposed to apply to different media formats, such as text, film and pictures. However, studies generally focused mostly on textual narratives. Furthermore, to the best of my knowledge, no study directly investigated if violations of external realism leads to lower transportation.

Third, I investigated how conflicting information affects subjective understanding and aesthetic emotions. The literature mostly considered beneficial effects of additional information on the subjective understanding and aesthetic evaluation of artworks. However, fluency theories of art processing as well as some empirical evidence (Belke et al., 2010; Gerger & Leder, 2015) suggests that additional information could also have negative consequences on the viewer's art experience. This is the case when non-expert viewers are confronted with conflicting information making the artwork and the artist's intentions more difficult to understand. In addition to previous research I thereby investigated whether these negative effects can be compensated by additional information explaining the conflict. According to the VIMAP the viewers encountering conflicts should experience certain emotions as a result of the schema congruency check detecting incongruence (Pelowski et al., 2017).

I investigated painting documents together with audio-texts in four studies (see table 1 for an overview). In Study 1, I manipulated the naming or not naming of discrepancies between the content of the painting and reality, as it can be assumed based on historians' opinions. The information about the source was presented as an additional text label together with each painting. I expected that recipients detect conflicts between text and picture documents indicated by longer processing of pictorial information when it is named with discrepancy than without discrepancy. In addition, if recipients engage in sourcing, processing times of source information should be longer when discrepancies are named than when they are not named. Alternatively, if recipients engage in information seeking, they

should become more interested in further information about the painting. Lastly, transportation should be lower with discrepancies named than without due to the less fluent construction of the mental model about the story.

In Study 2, I manipulated the naming or not naming of discrepancies again. In addition, audio-text subsequently either explained the conflict with benevolent intentions of the source or did provide information about the source not explaining the conflict. I expected that recipients consider the content together with information about the source. Therefore, the naming of discrepancies should lower the trustworthiness of the document when the discrepancy is not explained, but an explanation of benevolent intentions of the source should compensate this effect. In addition, I expected that content of the picture is better memorized when presented with discrepancies named than without. Lastly, I aimed at replicating the effect of a lower liking when discrepancies are named than when they are not named.

In Study 3, I used the same manipulation of discrepancy and explanation as in Study 2, but I focused on the aesthetic evaluation. For this reason, participants were framed in the instruction that the study is about artworks and comparable to a visit in a museum of art or art-history. In Study 3 I expected that the naming of discrepancies lowers the subjective understanding and liking of the paintings when no explanation is given, but an explanation should compensate these effects. I expected that the naming of discrepancies elicits more surprise than not naming the discrepancies. I expected that interest is lower when discrepancies are named but no explanation is given, and that an explanation should compensate this effect. In contrast, confusion should be higher when discrepancies are named but no explanation is given, and that an explanation should again compensate this effect. Lastly, again I aimed at replicating the effect of a lower liking when discrepancies are named than when they are not named.

In Study 4, the audio-text always named discrepancies between two picture documents depicting the same historic event. I manipulated the information about the source

by either presenting the intentions of both artists without difference, with difference, or not naming any intentions. When the artists' intentions were described without difference, both were described to intend propagandistic purposes. When the artists' intentions were described with difference one was described intending a documentary purpose and the other as intending a propagandistic purpose. When no intentions were described a text in equal length and similar in wording was presented not revealing any intentions. The information about the sources was presented prior to the paintings and could be viewed again during a task after the presentation of the paintings. I expected that with differences in intentions compared to the two other groups taken together, the information about the artists should be clicked more often during the task and should be processed longer. I expected that, during the task, the difference in processing times of the two paintings should be higher with differences compared to the two other groups since viewers should focus on the more trustworthy picture document and tendentially ignore the other picture document. I also expected that the viewers' trustworthiness rating of the content differs more greatly between the two paintings when differences in the group where differences in intentions are presented than in the two other groups. With differences in intentions compared to the other two groups, participants should use the source information more often indicated by citations in an essay, they should remember the source information better but should experience a lower interest in further information about the artist or the conflict since they can use the available source information to reconcile the conflict while participants in the other groups cannot.

In sum, in my four studies I systematically investigated how conflicting information affects three relevant aspects of pictures depicting historic events (see table 1). The main focus was on the picture as a document. Nevertheless, I considered the narrative and aesthetic character as an important part of picture processing especially in the case of history paintings. In the next chapters, I describe the studies I conducted. The studies are followed by a general discussion of the results. The chapters presenting the empirical studies can stand for their own and are written as separately readable manuscripts. For this reason,

the content from the empirical chapters can overlap with each other and with the content of this introduction and my general discussion.

Table 1

Overview of the Present Studies

Study	Topic	Independent Variables	Dependent Variables	Conflict	*N*
1*	historical document,	discrepancy (not-named vs. named)	fixation time on content, fixation time on source, transportation, interest	picture – audio text referring to “ reality”	63
2*	historical document, narration	discrepancy (not-named vs. named), Explanation (without vs. with)	trustworthiness, memory of content, transportation	picture – audio text referring to “ reality”	110
3**	artwork, narration	discrepancy (not-named vs. named), Explanation (without vs. with)	subjective understanding, liking, surprise, interest, confusion	picture – audio text referring to “ reality”	139
4	historical document	benevolence of two Artists (without differences, with differences vs. not mentioned)	clicks on source, processing time for source, processing time for content, trustworthiness, citations in essay, source memory, interest	picture – picture	129

*Study 1 and 2 are included in a published manuscript (Knoos, Glaser, & Schwan, 2021). **Study 3 is included in a published manuscript (Knoos, Glaser, & Schwan, 2021).

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Declaration according to § 5 Abs. 2 No. 8 of the PhD regulations of the Faculty of Science

-Collaborative Publications-

The following chapter (Chapter 2) consists of a manuscript that is published and was co-authored by Manuela Glaser and Stephan Schwan. The proportional contributions to this manuscript are presented in the subsequent table.

Author	Author position	Scientific ideas (%)	Data generation (%)	Analysis and Interpretation (%)	Paper writing (%)
Manuel Knoos	First Author	80	100	80	80
Manuela Glaser	Second Author	10	0	10	10
Stephan Schwan	Third Author	10	0	10	10
Title of the paper	Multiple Documents of Text and Picture: Naming a Historical Painting's Inaccuracies Influences Conflict Regulation Strategies				
Status in publication process:	Published. Knoos, M., Glaser, M., & Schwan, S. (2021). Multiple Documents of Text and Picture: Naming a Historical Painting's Inaccuracies Influences Conflict Regulation Strategies. *Contemporary Educational Psychology*, 101970. https://doi.org/10.1016/j.cedpsych.2021.101970				

2. Study 1 and 2

Multiple Documents of Text and Picture: Naming a Historical Painting's Inaccuracies Influences Conflict Regulation Strategies

In his seminal study on the use of multiple documents for historical thinking and problem solving, Wineburg (1991) presented eight different texts about the battle of Lexington together with three pictorial representations, asking both expert historians and students to give an account of the historical event addressed in the documents. Subsequent research has dealt extensively with the question of how readers integrate a set of written documents into a coherent mental representation, taking into account divergent perspectives and incongruencies between different text sources (Perfetti et al., 1999; Stadtler et al., 2014). In contrast, despite their use in the original Wineburg (1991) study, the role of pictorial representations and their interplay with texts for constructing a mental model of the historical event has remained largely unexplored. The present study aims to fill this research gap by investigating how recipients process combinations of historical paintings and simultaneously presented verbal descriptions addressing the historical painting and its inaccuracies when compared to today's most plausible version of the historic event.

Paintings as Historical Sources

In order to comprehend a historic event, pictures as well as text can be used as historical sources (Burke, 2001; Smith et al., 2019; van Boxtel & van Drie, 2012). Pictures are useful sources because they can record acts of eye-witnessing, such as the paintings by Louis-François Lejeune who produced a series of battle pictures during the Napoleonic Wars in which he participated. However, information about the event cannot always be easily derived by looking at a picture's content. For pictures, the same problems of context (van Boxtel & van Drie, 2012), function, rhetoric, and so on, emerge as they do for text (Burke, 2001). Therefore, the argument that content and author of a text must be considered together to comprehend a document's validity (Wineburg, 1991) is equally true for pictures of historic events. Painters in the same manner as authors often distort the historical facts

presented in their work (Burke, 2001) and especially non-expert viewers might easily overlook these distortions. For this reason, viewers might benefit from verbal information such as labels and captions, audio and personal guides in museums, or by teachers' oral explications and written text in history textbooks. The picture thereby provides information about the historical event; the label or caption gives information about the painting's artist and social context, and the accompanying verbal information may refer to the painting, other sources of the historic event as well as the artist. Moreover, in the verbal information, the painting's discrepancies between the depicted narrative and today's most plausible version of the historic event can be explicitly named and explained. Consequently, both in museum and school situations, viewers face the task of integrating information from these multiple, partly conflicting, documents (Britt & Rouet, 2012).

Processing and Resolving Conflicting Information from Multiple Text Documents

According to the conflict source integration model (CSI; Stadtler & Bromme, 2014), conflicting information is processed in three stages. First, readers must detect conflicting claims. This is followed by the reader's conflict regulation, as a second step, which may or may not end with conflict resolution, as the final step.

To detect conflicts, readers need to co-activate the conflicting propositions in working memory. When a conflict is detected, readers observe a breakdown in text coherence leading to longer reading times for conflicting compared to non-conflicting information, both in a single text (Albrecht & O'Brien, 1993; Kendeou et al., 2013; O'Brien et al., 1998; Rapp & Kendeou, 2007) and across multiple texts (Beker et al., 2016). Furthermore, prolonged reading times can result in better memory of the conflicting information (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995).

Conflict regulation can take various forms. First, readers can reconcile the conflict by drawing inferences to explain the conflict (Blanc et al., 2008). Second, readers may ignore the conflict, which means that they detect the conflict but do not remember it afterwards. This tendency is fostered when the conflict seems irrelevant to the reader (Stadtler et al., 2012).

In this case, it can be argued that the reader did not engage in a regulatory strategy (Stadtler & Bromme, 2014). However, readers who do not recall the conflict might have engaged in a regulation strategy by distorting parts of the conflicting information and, therefore, instead of not remembering the information, they remember the conflicting information wrongly as being consistent (Hakala & O'Brien, 1995).

Third, if readers cannot easily solve or fail to solve the conflicts due to insufficient information, they might postpone the conflict resolution and become more interested in receiving further information (Graesser et al., 1994; Stein & Trabasso, 1985). This is supported by studies showing that readers encountering conflicting information ask more information-seeking questions than readers not confronted with a conflict (Graesser & McMahen, 1993; Otero & Graesser, 2001). The relation between encountering conflicts and subsequent information-seeking may be grounded in the perception of knowledge gaps and the motivation to fill them. Readers encountering conflicts (Berlyne, 1960) and perceiving knowledge gaps (Loewenstein, 1994) develop motivational states, such as interest or curiosity, to acquire information in order to reduce uncertainty and close the knowledge gaps. This is in line with the reward-learning framework (Murayama et al., 2019), postulating that perceived knowledge gaps are the key reason for learners to engage in information seeking behavior, especially when they expect a feeling of reward by filling their knowledge gaps.

Fourth, according to the CSI model (Stadtler & Bromme, 2014), readers can also accept the information as being conflicting due to the different sources. Thereby they may also try to solve the discrepancies by asking what is true or whom to believe. This is supported by evidence in the context of the discrepancy-induced source comprehension model (D-ISC; Braasch & Bråten, 2017). Readers of texts with conflicting sources spend more effort processing author information and have a better memory for links between authors and contents afterwards than readers of texts without conflicting sources. They thereby evaluate the author information and consider the information that allows them to judge the reliability or trustworthiness of the document (cf. Saux et al., 2018). To explain the

discrepancies, they can use the available information about the authors, such as their expertise or benevolence (Thomm & Bromme, 2016). Benevolence is the recipient's perception of how much the author intends something good for the recipients and whether she or he provides the best possible information instead of following selfish goals (Mascaro & Sperber, 2009; Thomm & Bromme, 2016). It was shown that discrepant information especially enhances the memory for features of the author that help to explain the discrepancy, such as expert knowledge, but not for source features that are irrelevant for an explanation, such as the appearance of an author (Saux et al., 2018). Alike for benevolence, it was demonstrated that source information is processed longer when two conflicting sources differ in their benevolence than if they do not differ in benevolence. This experiment showed that readers did indeed use the information about benevolence when it was helpful for explaining the discrepancy (Gottschling et al., 2019).

The Role of Pictures in Processing Multiple Documents

Several studies directly compared multiple document situations with either only texts, only videos, or text and videos (List, 2018; List & Ballenger, 2019; Salmerón et al., 2020). Mixed results were found for comprehension strategies used by the learners across different types of media. One study found that learners mostly used the same strategies to comprehend information represented in two texts versus in two videos and had a similar good comprehension about the topic (List, 2018). Another study (List & Ballenger, 2019), on the other hand, showed that learners using sequentially presented texts compared to sequentially presented video or sequentially presented text and video material reported higher use of comprehension strategies, such as cross-textual elaboration. In addition, comprehension of the first document was better when information was presented in a text than in a video. However, in these studies no conflict between the documents was present and only general comprehension strategies, but not specific conflict regulation strategies, were examined.

One study used conflicting documents in the form of only texts, only videos, or text and video (Salmerón et al., 2020). The results revealed that integration was better with texts than videos, but learners defended the opinion presented in the video more often than when the opinion was presented in the text. Overall source memory was not affected by the different presentation modes, but learners better remembered the profession of one author if presented by visual cues in the videos than by verbal statements in texts. This study therefore indicates that not only conflicts influence sourcing and source memory but also how sources are presented by the possibilities of different media types. Thus, design principles known from multimedia research (Mayer & Fiorella, 2014) could be useful as recently investigated by a study using text and video documents (Merkt & Huff, 2020). This study manipulated the spatial and temporal position of the information about the authors (spatially and temporally integrated with document contents vs. spatially and temporally separated from document contents) and the activation of schemata (by presenting author information previous to vs. after the documents) in a multiple document situation. Based on the cognitive theory of multimedia learning the authors hypothesized better performance in a retention test asking participants about differences between the documents when information and source were integrated compared to separated. They further hypothesized better performance when presenting source information before than after the content of the documents. The results of Experiment 1 supported neither of the hypothesized main effects but an interaction. The performance was better when source information was integrated and presented before the contents than after the contents but when source information was separated, performance was better when source information was presented after the contents. The pattern of results is complex but indicates that considering principles from other fields such as multimedia research can be a fruitful extension of multiple document research especially when considering documents in different media formats.

In videos, the source might additionally be made visible in the video and not just described by text, which could affect source memory (Salmerón et al., 2020) due to a potential dual-coded representation of the source information (Paivio, 1990). Information

presented in videos is more transient than written text and the same applies to audio-texts (Wong et al., 2012), especially if the information cannot be viewed or heard again. Thereby conflicts as well as information about the source might be missed more easily by the recipients. However, with audio texts accompanying a painting, the conflicting information can be presented simultaneously instead of sequentially, leading to a higher temporal contiguity (Ginns, 2006), which could foster the detection and memory of conflicts. Therefore, investigation of documents in other media formats than texts or in mixed sets of texts, pictures, and videos are of relevance for multiple document research.

Conflicts between textual and pictorial information in multimedia learning material were investigated in two studies by Schüler (2017, 2019). In these experiments, eye movements instead of reading times were recorded. In the first experiment by Schüler (2017), pictures of maps were accompanied by consistent written text describing the different elements of the map. The last element of the text was either consistent or inconsistent with the picture, for example, regarding the location or color of a pictorial element. Compared to consistent information, inconsistencies resulted in longer fixation times on the text and picture, but less pictorial content was recalled afterwards. In the second experiment by Schüler (2017), the mechanism of a toilet flush was used as material. Again, fixation times were longer in the inconsistent condition, but only for the text and not for the picture, while no significant differences for the learning outcomes were observed. Additionally, about four-fifths of the participants did not remember the conflict and therefore may have ignored it or distorted parts of the conflict in order to restore coherence. Another study used multiple documents of either only text or of text and picture, explaining the formation of a tornado (Schüler, 2019). The results revealed that fixation counts were higher on the text and on the picture when these were inconsistent with previously presented textual and pictorial information compared to when they were consistent. Again, there was no difference regarding the learning outcomes between the group receiving inconsistent information and the group receiving consistent information, indicating that both groups were able to construct a coherent mental model. It seems plausible that because the learning material was

presented twice, nearly two-thirds of the participants who received the conflicting information remembered it. For this reason, the participants may have not ignored the conflict but rather inferred an explanation or concluded that parts of the information were wrong.

In sum, the results by Schüler (2019) indicate a fundamental similarity in the processing of discrepancies between multiple texts and text-picture-combinations that may also generalize to multiple documents of other media formats: longer processing times as an indicator of co-activation of information and conflict detection. However, besides this similarity, research comparing multiple texts with multiple videos (List, 2018; List & Ballenger, 2019) showed mixed results with regard to comprehension strategies and with regard to outcomes of conflict regulation comparing processing of multiple text, multiple videos, and text-video-combinations the study by Salmerón et al. (2020) showed different outcomes for retention of source information between these conditions. Hence, comprehension and conflict regulation strategies leading to this different outcome may have been affected differently by aspects resulting from different media formats such as transiency or the possibility to present multiple documents simultaneously.

At present, pictures as well as text-picture combinations are widely ignored in multiple document research, although they can be found in many authentic multiple document situations such as newspapers, social media, advertisement, or museums. In addition, in many authentic situations, the documents explicitly address discrepancies with other sources on the same topic (Braasch & Bråten, 2017), whereas explicit naming of discrepancies, to the best of our knowledge, has not been well examined in multiple document research. Therefore, it is still an open question whether learners apply similar strategies for conflict regulation, such as sourcing, for processing multiple text documents as for processing multiple pictures or multiple documents of mixed media formats. By addressing the case of a simultaneous presentation of pictorial documents together with auditory commentaries as a second document class, the present studies aim to provide some insights into this topic.

Transportation into the Narrative Depicted in a Historical Painting

With different media formats and mixed media presentations, the immersion potential of the different documents should be considered as a factor that influences their processing, especially when their contents are inherently narrative as it is the case with documents about historical events. Using conflict regulation strategies to resolve discrepancies between a verbal description and a pictorial depiction of historical events may, as a side effect, make the viewer's construction of the mental model of these narratives less fluent. According to the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008), this may lead to a reduction in perceived transportation. Transportation is defined as the reader's or viewer's experience of being mentally absorbed in the story world (Gerrig, 1993). It consists of a cognitive, an emotional, and an imaginative component (Green and Brock, 2002). Transported individuals focus their cognitive processing on the events described in the narrative, identify and feel with the characters, and create vivid mental images of the places and characters. They show participatory responses (Gerrig, 1993), lose awareness for what is going on in the real world around them, may also become less critical of information presented in the narrative (Green & Brock, 2000), and less likely to detect discrepancies (Busselle & Bilandzic, 2008).

According to Busselle and Bilandzic (2008), the default mode of processing narratives is to believe them to be true (see also Gilbert, 1991). Therefore, readers or viewers usually construct a mental representation of a narrative without taking the epistemological status of it into account. Only in cases in which the fluent generation of a mental model of the narrative is disturbed, does an effortful analytical and critical processing of the narrative take place. Readers or viewers, for example, do not evaluate the narrative's realism as long as no unexplained discrepancies between the story and the real world or within the narrative are encountered. When such a discrepancy is observed, the process of generating a mental model of the narrative and therefore also transportation is hampered, critical thinking is prompted, and individuals disengage from the narrative. In this case, readers or viewers switch from an involved narrative processing mode into a more distant, analytical processing mode (Vorderer, 1992) in which they are aware that the narrative has been created and

critically evaluate it. The default mode of fluently reconstructing the narrative and the effortful mode of critical processing, including effortful endeavors to solve the conflict and the possibly integration of the conflicting content into a mental model, can therefore be regarded as mutually exclusive. Nevertheless, in the end, both lines of processing result in the integration of information and the generation of a coherent mental model, provided that encountered inconsistencies are successfully solved by subsequent critical, analytical processing. However, the latter should be accompanied by an increased invested effort and by less transportation.

Although models of narrative processing and transportation have been examined mainly using written texts and movies, they are assumed to apply to all forms of media, such as text, videos, and pictures (Green & Brock, 2000). For the present studies, we therefore assume that paintings depicting historical events (and thereby inherently present narratives) will also be processed in a narrative mode, resulting in a mental model of the historical event. If this occurs fluently, high levels of transportation into the historical narrative should be experienced. However, if conflicting information is detected during this narrative processing, transportation should be hampered. As a result, the recipient will feel no longer present in the historical event but will tend to experience it from an outside perspective. In order to restore the coherence of the historical narrative, the recipients may engage in different conflict regulation strategies, as discussed previously. If the recipients perceive a knowledge gap being responsible for the inconsistency, interest may be triggered to gain further information to close the gap, solve the conflict, and restore coherence by information seeking behavior (Berlyne. 1960; Loewenstein, 1994). Hence, while transportation implies being involved in the fluent processing of a narrative, interest in further information occurs when this fluent processing is no longer possible and implies being involved in solving conflicts to restore narrative coherence. Transportation and interest in further information, therefore, are related to different aspects of processing a narrative. This difference has to be taken into account when examining the influence of inconsistencies between multiple documents of mixed

media formats on conflict regulation strategies and the experience of transportation in the present studies.

Experimental Overview

In summary, empirical evidence suggests longer processing of discrepant content with the naming of discrepancies than without, which could also lead to better memory for the discrepant content (e.g. Albrecht & O'Brien, 1993). According to the literature on multiple documents, readers use conflict-regulation strategies when they are confronted with conflicting information (Stadtler & Bromme, 2014), resulting in longer processing of source information and the use of source information to explain the conflict (e.g. Gottschling et al., 2019), or in postponing the conflict resolution and becoming more interested in further information (Stein & Trabasso, 1985). The model of narrative engagement (Busselle & Bilandzic, 2008) suggests that discrepancies lead to a lower experienced transportation due to the hampered construction of the story's mental model. We conducted two experiments to examine conflict regulation strategies and experienced transportation with multiple documents of mixed media formats (pictures and audio-texts). This situation differs from multiple documents of only text in several ways. First, the information from the picture is not necessarily processed linearly, whereas the processing of texts is linear, with Western readers reading texts from left to right and top to bottom. Pictures depict stories more vividly than texts (Burke, 2001) and are therefore more immersive than texts. This raises the importance of considering aspects of information processing such as transportation especially when considering multiple documents in different media formats. The information presented in audio-texts about the conflict and the authors is more transient than written text information (Wong et al., 2012), which can be processed at one's own pace. However, an audio-text can be presented simultaneously to a picture while two or more texts must be processed sequentially. We were interested whether the claims of the CSI generalize from studies using only text documents to such a different multiple document situation including mixed media formats.

In Experiment 1, we examined the effects of naming or not naming a painting's discrepancies on attention to the painting's content and its source label and the effects on interest as indicators of conflict-regulation strategies. In addition, we investigated the effects of discrepancies on transportation. We hypothesized that with discrepancies named the viewers will allocate more attention to the pictorial element and to the source label . Viewers will also report higher interest but experience lower transportation with discrepancies named compared to without.

In Experiment 2, we investigated the effects of naming or not naming discrepancies, combined with either an explanation of the discrepancies (benevolent intention of the artist) or not, on the evaluation of the trustworthiness of the historical painting and the retention of the manipulated pictorial elements. We hypothesized that viewers rate the trustworthiness of the painting lower with discrepancies named than without and that this effect is compensated by a benevolent explanation about the authors intentions for the discrepancies. We also hypothesized higher retention of manipulated pictorial elements with discrepancies named than without. Lastly, in Experiment 2, we aimed to replicate the effect on transportation found in Experiment 1.

Experiment 1

Experiment 1 focused on the distribution of attention as indicated by fixation times on the discrepant pictorial elements and the source label. Additionally, it focused on interest in further information as an indicator of different top-down conflict regulation strategies as well as on transportation as an indicator of the fluent generation of the mental model of the narrative depicted in the painting. Participants were provided with the multiple sources typical for learning history in formal (e.g. schoolbooks) and informal (e.g. museums) learning settings, namely, a picture of a painting, a label with the painting's title, the name of artist, the year of production, how much time had passed between the event and the production of the painting, and a spoken text describing the historical event and commenting on its portrayal in the painting. The spoken texts either explicitly named discrepancies between the depicted

narrative of a painting and today's most plausible historical narrative (discrepancies named condition) or did not address these discrepancies (discrepancies not named condition). We expected that naming discrepancies in the audio-texts will lead to closer inspection of discrepant elements in the picture and conflict regulation strategies, like paying attention to source information about the painter, or a need for further information.

If viewers recognize the conflicts, they should fixate the pictorial elements longer in the condition with discrepancies explicitly named compared to discrepancies not named (H1). In particular, if viewers explain the conflict due to different source qualities, the label with the information about the artist should be fixated longer in the condition with discrepancies explicitly named compared to discrepancies not named (H2). Further, if viewers postpone the conflict regulation or cannot fully explain the conflict, the interest in additional information about the painting should be higher with than without discrepancies explicitly named (H3). Because both the strategy of sourcing indicated by longer fixation times on the label (H2) and the postponing of conflict regulation indicated by interest in further information (H3) are based on searching for additional information to resolve the conflict, they are assumed to occur alternatively. The longer fixation times on the pictorial elements (H1), on the other hand, can foremost be seen as an indicator of the bottom-up co-activation of conflicting information. Therefore, they are considered to be an indicator of conflict recognition. Therefore, Hypothesis 1 is regarded as being complementary to Hypotheses 2 and 3. In addition, the experience of transportation was measured as an indicator of narrative processing. We therefore assume that transportation should be lower with discrepancies named than without discrepancies named by an audio-text (H4). Finally, we exploratively looked at the retention of the pictorial elements.

Method

Participants

From our institute's mailing list, 94 participants were recruited. Fifteen were excluded because they were familiar with one or more paintings used in the study. Another five

participants were excluded due to problems with their eye-recordings. Three of these participants had an inaccurate recording with a deviation of more than 0.80°; one had a tracking ratio of lower than 80%, and for one, the data were not recorded by the device. Ten participants did not follow the instructions; seven of these did not listen to the whole audio-text, and three answered the free recall for the wrong painting. One was excluded because of not being a native German speaker. Ultimately, 63 participants remained for the analysis (55 female, 8 males; aged between 18 and 32 years, M = 22.97, SD = 3.02). They were from a broad range of disciplines (e.g. psychology, law, medicine, economics, natural sciences) and were randomly assigned to one of our two conditions (n = 31 with discrepancies named, n = 32 with discrepancies not named).

Design

Our hypotheses with regard to fixation times, interest, and transportation were tested in a mixed 2x6 design with naming of discrepancy (with vs. without) as the between-subjects factor and painting (Dong, Manet, Gros, Hellqvist, Leutze, West) as the within-subjects factor. For the exploratory research question regarding retention, a free recall test was conducted for the last painting presented (either Leutze or West). For the analysis of retention, painting was included as a between-subjects factor, resulting in a 2 (with vs. without naming of discrepancy) x 2 (painting by Leutze vs. West) design.

Technical Equipment

The experiment was run with SMI Experiment Center 3.7.68, which was run on two computers with 23-inch Dell monitors [1920x1080px]. A 250 Hz remote eye-tracking system and the eye-tracking software IView RED 4.4 from the company Senso Motoric Instruments (SMI) were used. The eye-tracking cameras were mounted below the monitors. The participants were seated approximately 68 cm in front of the screen. Their chins were supported on a chin rest to control for constant eye-to-screen distance and head-movements. They navigated with the keyboard through the experiment. The audio-texts were

presented via a headset. The eye-tracking data analysis was performed with SMI BeGaze 3.7.59 with the program's default setting.

Material

Pictures of six historical paintings were chosen: "The Founding Ceremony of the Nation" (1967) by Dong Xiwen, which is exhibited in the National Museum of China in Beijing, "The Execution of Emperor Maximillian of Mexico" (1869) by Edouard Manet, exhibited in the Kunsthalle in Mannheim, "Napoleon Visits the Plague Victims of Jaffa" (1804) by Antoine-Jean Gros, exhibited in the Louvre in Paris, "Valdemar Atterdag Holding Visby to Ransom" (1882) by Carl Gustaf Hellqvist, exhibited in the Nationalmuseum in Stockholm, "The Death of General Wolfe" (1770) by Benjamin West, exhibited in the National Gallery of Canada in Ottawa, and "Washington Crossing the Delaware" (1851) by Emanuel Leutze, exhibited in the Metropolitan Museum of Art in New York. A short label of written text about the painting's artist, his nationality, and the time when he created the painting was placed beside the lower left side of the respective picture (see Appendix A). We used "The Execution of Lady Jane Grey" (1833) by Paul Delaroche, which is exhibited in the National Gallery in London, and its label as an example in the introduction to the experimental procedure, while the other pictures of the paintings and corresponding labels were used for the analyses.

For each painting, two audio-text versions were created that differed according to the experimental condition in naming versus not naming discrepancies between today's most plausible version of the historic event and its depiction in the painting. Both versions consisted of six segments. The participants continued self-paced from segment to segment by pressing the space bar. The first segment was a short introduction about the painting and the historical event to which the painting refers. The second segment described the main character of the depicted narration. In the next four segments, single pictorial elements were described. The description of each pictorial element had a three-part structure of localization (where is it located in the painting), description (what does it look like), and interpretation (how is it related to the historical narrative). While the third and fourth segment each

described one non-discrepant pictorial element, the fifth and the sixth segment each described one pictorial element for which the verbal explanation was varied. The two conditions differed with regard to the last sentence of the interpretations in the fifth and sixth audio-text segments. Here the respective pictorial elements were either named as being discrepant to today's most plausible account of the historical narrative (experimental condition) or not (control condition). The audio-texts of the two conditions were equally long with only a few words changed. For example, the last sentence for the sun depicted in the painting of Leutze was (1) without naming the discrepancy: "However, the crossing was carried out and finished during the daytime." (2) and with naming the discrepancy: "However, the crossing was carried out and finished during the night." The audio-texts of the different paintings varied between 1:44 and 2:08 minutes (see Appendix B for an example text).

Measures

Manipulation Check. The naming of discrepancies between a historical painting and the historical evidence should reduce the perceived realism of the painting. In order to check whether the participants in the experimental condition actually perceived the paintings as being less realistic than the participants in the control condition, the participants were asked to rate with one item how realistic the depicted event appeared to be to them. The answer had to be given on a 7-point Likert scale ranging from one ("very little realistic") to seven ("very much realistic").

Fixation Times. For each of the six paintings, an area of interest (AOI) was defined for each of the two pictorial elements that were either named as being discrepant or being not discrepant to the historical evidence. A third AOI was defined for the text label of each painting. The fixation times on each AOI were measured in milliseconds. A fixation was defined as an eye movement at a speed less than 40°/s, which is the default setting of the software. The recordings of the fixation times on the AOIs of the pictorial elements and the label started from the naming of the element as either discrepant or not discrepant until the participants continued self-paced to the next segment of the audio-text. Mean scores for the

fixation time on the text label AOIs and the AOIs of the two manipulated pictorial elements were calculated for each painting.

Interest. Interest in more information about the painting was measured after each painting with a single item: “How much are you interested in receiving further information about the painting from the audio-text?”. Answers were given on a 7-point Likert scale ranging from one (“very little”) to seven (“very much”).

Transportation. Transportation into the historic narrative was measured after each painting by a single item: “How much did you feel yourself transported into the historic event by the painting?” The answers had to be given on a 7-point Likert scale ranging from one (“very little”) to seven (“very much”).

Retention. Retention of the pictorial elements of the paintings was measured using a free-recall test (see Appendix D), which was adapted from Glaser and Schwan (2015). The participants were asked to write a description and draw a sketch of the last painting presented (either Leutze or West, counterbalanced within the conditions). Testing time was limited to ten minutes. The verbal description and the drawing were considered complementary and were rated with regard to the correct retention of the 16 predefined meaningful pictorial elements for each painting. The participants received one point for each correctly remembered pictorial element, and the mean score was calculated. The recall was rated by two independent raters with an interrater reliability (Krippendorff’s alpha) of $\alpha = .86$. In cases in which these raters did not agree, a third rater made the decision.

Procedure

Up to two participants were tested in each session. First, they were welcomed and seated in front of the computer screen. The eye-tracking device was adjusted based on a 9-point calibration and a 5-point validation. The participants received the instruction that they will view seven paintings depicting historical events together with an audio-text accompanying each painting. The instruction explained how to navigate through the audio

segments by pressing the space bar and that it was not possible to hear an audio segment again. The participants were informed that they will have to answer questions about the paintings after viewing them and therefore should pay close attention to the paintings and the audio-text. After the instruction about the experiment, the pictures of the paintings with the source label about the artist were presented sequentially together with their respective audio-texts. Each painting was introduced by a written introduction stating in German: “In the following, you will see the painting [title of the painting] by the artist [name of the artist].” After this, the presentation of a fixation cross followed, which had to be fixated for two seconds to proceed automatically to the presentation of the painting. The first painting by Delaroche depicted “The Execution of Lady Jane Grey”, and was followed by the six test paintings: “The Founding Ceremony” by Dong, “Napoleon Visits the Plague-Stricken” by Gros, “Valdemar Atterdag Holding Visby to Ransom” by Hellqvist, “The Execution of Emperor Maximilian” by Manet, “Washington Crossing the Delaware” by Leutze, and “The Death of General Wolfe” by West. The presentation sequence of Leutze and West changed within both conditions so that for one half of the participants, the last presented painting was the Leutze painting, and for the other half, the last painting was the West painting. This was done in order to vary the painting for the retention about the last painting.

After starting the audio-visual presentation of a painting, the participants navigated self-paced through the presentation. At the beginning of the presentation, they were free to read the label and look at the painting for as long as they wanted without audio information. The first segment of the audio-text began when the participants pressed the spacebar and they continued after one segment to the next segment by pressing the spacebar again. When the audio presentation was done, the painting with its label turned off, the manipulation check, and the measures about transportation and interest were presented. After all audiovisual presentations and the measurements of interest and transportation, a filler task of about seven minutes followed, consisting of a verbal and pictorial memory puzzle presented in order to inhibit further memorization of the previously seen information. After the puzzle, the retention test was handed out to the participants (either about the Leutze or the

West painting depending on the last painting they had seen in the audiovisual presentation). The participants had 10 minutes to answer the retention test. Finally, demographic variables (age, gender, and profession) were collected. In addition, they were asked whether they had already known any of the paintings that were presented to them previously and if yes which ones. The participants were debriefed and paid seven Euros for participation. The whole experiment lasted approximately 45-50 minutes. The study received institutional research ethics committee approval.

Results

Manipulation Check: Perceived Realism of the Paintings

The ANOVA for perceived realism revealed a significant main effect of discrepancy, $F(1, 61) = 21.11$, $p < .001$, $\eta^2_p = .257$. Perceived realism was rated significantly lower with discrepancies named compared to without discrepancies named (see Table 2). The main effect of painting was significant, $F(5, 305) = 14.08$, $p < .001$, $\eta^2_p = .188$. Bonferroni-adjusted comparisons showed that the Dong painting was rated as more realistic than Leutze, Hellqvist, Manet, Gros, and West (see Table 3). The interaction between discrepancy and painting was significant, $F(5, 305) = 3.66$, $p = .003$, $\eta^2_p = .057$. Five of six paintings were rated as less realistic with discrepancies named compared to without discrepancies named: Dong ($M = 3.90$, $SD = 1.22$ vs. $M = 5.34$, $SD = 1.10$, $p < .001$), Manet ($M = 2.61$, $SD = 1.20$ vs. $M = 4.03$, $SD = 1.45$, $p < .001$), Hellqvist ($M = 2.77$, $SD = 1.33$ vs. $M = 4.13$, $SD = 1.43$, $p < .001$), Gros ($M = 2.71$, $SD = 1.37$ vs. $M = 3.75$, $SD = 1.46$, $p = .005$), and West ($M = 2.90$, $SD = 1.19$ vs. $M = 3.78$, $SD = 1.45$, $p = .011$), but for the Leutze painting there was no significant difference between the conditions with ($M = 3.48$, $SD = 1.15$) and without discrepancies named ($M = 3.53$, $SD = 1.62$), $p = .895$.

Table 2

Descriptive Data (Means and Standard Deviations) for the Dependent Measures of Experiment 1 by Condition

	Without discrepancy named	With discrepancy named
Perceived realism	4.09 (0.92)	3.06 (0.85)
Viewing times[1]	20362.80 (1175.47)	20409.68 (834.84)
Fixation label[1]	234.07 (390.27)	159.77 (202.54)
Fixation picture[1]	1277.79 (556.70)	1530.74 (332.03)
Interest	4.37 (0.71)	3.99 (0.86)
Transportation	4.05 (0.81)	3.63 (0.73)
Retention of picture	0.35 (0.09)	0.37 (0.10)

Note. [1]Fixation times and viewing times in milliseconds

Viewing Times at Audio-Text Sections in which the Discrepancies were either Named or not Named

As the audio-visual presentation was self-paced, we checked for general differences in self-paced viewing times at audio-text sections in which the discrepancies were either named or not named. The ANOVA of the viewing times at these sections revealed no significant main effect of discrepancy, $F(1, 61) = 0.03$, $p = .856$, $\eta^2_p = .001$. The main effect of painting was significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(4.04, 246.24) = 49.26$, $p < 001$, $\eta^2_p = .447$, with longest viewing times for the Manet painting and shortest viewing times for West painting. The interaction between

discrepancy and painting was not significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(4.04, 246.24) = 1.43$, $p = .226$, $\eta^2_p = .023$. The analysis of viewing times shows that potential differences between the conditions with regard to our dependent variables are not attributable to different self-paced viewing times.

Table 3

Descriptive Data (Means and Standard Deviations) for the Dependent Measures of Experiment 1 by Painting

	Dong	Leutze	Hellqvist	Manet	Gros	West
Perceived realism	4.63	3.51	3.46	3.33	3.24	3.35
	(1.36)	(1.40)	(1.53)	(1.50)	(1.50)	(1.39)
Viewing times[1]	20816.28	19642.66	20810.77	21154.52	20344.79	19546.21
	(1460.23)	(996.68)	(1545.92)	(889.76)	(1249.41)	(1288.31)
Fixation label[1]	164.88	106.29	198.52	524.50	73.64	117.20
	(459.13)	(369.76)	(489.94)	(1311.68)	(288.75)	(339.65)
Fixation picture[1]	1734.23	807.60	1023.29	2542.93	1114.56	1190.92
	(911.12)	(650.61)	(719.78)	(834.59)	(665.72)	(1018.80)
Interest	4.10	4.10	4.27	3.95	4.68	4.00
	(1.35)	(1.33)	(1.39)	(1.34)	(1.09)	(1.22)
Transportation	4.08	3.81	4.06	3.41	4.06	3.63
	(1.30)	(1.18)	(1.26)	(1.32)	(1.31)	(1.20)

Note. [1]Fixation times and viewing times in milliseconds

Fixation times on discrepant pictorial elements

The ANOVA of fixation times for discrepant pictorial elements at audio-text sections in which the discrepancies were either named or not named revealed a significant main effect of discrepancy, $F(1, 61) = 4.65$, $p = .035$, $\eta^2_p = .071$. The participants in the condition with discrepancies named fixated the discrepant pictorial elements significantly longer than the participants without discrepancies named. The main effect of painting was significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(4.07, 248.12) = 50.39$, $p < .001$, $\eta^2_p = .452$, with longest fixation times for discrepant elements of the Manet painting and shortest fixation times for discrepant elements of the Leutze painting. The interaction between discrepancy and painting with regard to fixation times on the manipulated pictorial elements was not significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(4.07, 248.12) = 1.65$, $p = .160$, $\eta^2_p = .026$.

Fixation Times on Source Labels

The ANOVA of fixation times on the source label revealed no significant main effect of discrepancy, $F(1, 61) = 0.89$, $p = .349$, $\eta^2_p = .014$. The main effect of painting was significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(1.85, 112.88) = 4.45$, $p = .016$, $\eta^2_p = .068$; however, Bonferroni-adjusted comparisons revealed no significant differences between the paintings. The interaction between discrepancy and painting with regard to fixation times on the source label was not significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(1.85, 112.88) = 0.37$, $p = .673$, $\eta^2_p = .006$.

Interest

The ANOVA for interest for additional information revealed no significant main effect of discrepancy, $F(1, 61) = 3.66$, $p = .060$, $\eta^2_p = .057$. The main effect of painting was significant, $F(5, 305) = 3.68$, $p = .003$, $\eta^2_p = .057$. The interaction between discrepancy and painting was not significant, $F(5, 305) = 0.41$, $p = .839$, $\eta^2_p = .007$.

Transportation

The ANOVA for transportation revealed a significant main effect of discrepancy, $F(1, 61) = 4.46$, $p = .039$, $\eta^2_p = .068$. The participants reported significantly lower transportation in the condition with discrepancies named than the participants in the condition without discrepancies named. The main effect of painting was significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(4.14, 252.49) = 4.16$, $p = .002$, $\eta^2_p = .064$, with the highest transportation reported for the Dong painting and the lowest transportation reported for the Manet painting. The interaction between discrepancy and painting was not significant using the Greenhouse-Geisser adjustment to correct for violations of sphericity, $F(4.14, 252.49) = 0.19$, $p = .946$, $\eta^2_p = .003$.

Retention of Pictorial Elements

We looked exploratively at the retention of the 16 pictorial elements of either Leutze or West. The ANOVA for the free recall test on retention of the pictorial elements revealed no significant main effect of discrepancy, $F(1, 59) = 1.31$, $p = .257$, $\eta^2_p = .022$. The main effect of painting was not significant, $F(1, 59) = 1.80$, $p = .184$, $\eta^2_p = .030$. The interaction between discrepancy and painting was not significant, $F(1, 59) = 0.67$, $p = .416$, $\eta^2_p = .011$.

Discussion

Experiment 1 investigated how the naming of historical inaccuracies in historical paintings influences the processing of and memory for information from a painting, a source label, and an audio-text. The results of the analysis of the fixation times on the content indicate that the viewers tended to allocate more visual attention to the discrepant pictorial content due to the recognition of conflicts (Hypothesis H1). The results neither support Hypothesis H2 that viewers regulated the conflict by allocating their attention to source information nor H3 that viewers postponed the conflict resolution and become more interested in receiving further information for solving the conflict. Furthermore, in support of H4, the viewers in the condition with discrepancies named reported significantly lower

transportation than those viewers in the condition without discrepancies named. With regard to the exploratory analysis of the retention of pictorial elements, no differences between the conditions with discrepancies and without discrepancies named were found. It is also noteworthy that 52% of the participants in the discrepancies named condition spontaneously reported the existence of discrepancies between the painting and the historical event in the free-recall task without being asked. This is in line with previous research indicating that explicit information about discrepancies enables viewers to tag the information as being discrepant (Butler et al., 2009).

Also, in line with previous studies, the longer fixation times on the discrepant content indicate that explicitly named discrepancies led participants to engage in bottom up-mechanisms of co-activation. The lower transportation with discrepancies named compared to without them named might indicate that viewers change from a narrative processing to a more critical, analytical processing when confronted with a conflict and thus reason about possible explanations for the conflict. However, we did not find support for any of the other proposed top-down conflict regulation strategies that are tightly linked to the co-activation of conflicting information (Stadtler & Bromme, 2014). More specifically, the viewers did not tend to ignore or distort parts of the conflict since the majority remembered the conflicting information afterwards. Also, explicit naming of discrepancies did not lead the viewers to postpone the conflict regulation, as this would have been indicated by a higher interest in receiving further information explaining the conflict. Lastly, the viewers also did not allocate more visual attention to source information provided via the label. Therefore, as the last option, the participants might have inferred an explanation.

Our results are partly compatible with previous studies of processing multiple conflicting text documents. On one hand, we also found a longer processing of content with discrepancies compared to without (Beker et al., 2016). On the other hand, and in contrast to the literature (Braasch & Bråten, 2017), source information was not processed longer in our experiment. In order to investigate more closely whether source information is used when

discrepancies between paintings and historical events are present, we varied the information about the source's benevolence in Experiment 2. This approach has already been similarly employed in multiple documents studies (Gottschling et al., 2019; Kammerer et al., 2016). We therefore asked the viewers to rate the trustworthiness of the painting's depiction of the historical event, depending on whether the discrepancies between painting and historical evidence were named or not and whether additional background information about the benevolent intentions of the artist was given or not.

Experiment 2

In Experiment 2, we investigated the effect of explicitly naming versus not naming discrepancies between a painting's depiction of a historical event and the historical evidence, combined with either an explanation of the discrepancy (benevolent intention of the artist) or not, on the evaluation of the trustworthiness of the historical painting, the retention of the manipulated pictorial elements, and the experienced transportation in the depicted narrative. The information about the artist's intention was given in the accompanying audio-text in a way that it either did or did not explain the discrepancy. This explanation always presented the respective artist as benevolent in such a way as to inform the participants that he had intended something good for the audience. For example, it was explained that West included a depiction of a Native American (who was not there in the historic event) to make the location of the event outside of Europe clearer to the European viewers.

In addition, Experiment 2 investigated the effects of named discrepancies on retention in more detail. Regarding the effects of discrepancies on retention, previous studies with multiple texts indicate a memory enhancement due to the deeper processing of discrepant information. These studies measured the recall of those bits of information that were presented either as discrepant or not (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995). In contrast, previous studies with picture and text did not find support for the enhancement of memory for pictorial content due to discrepancies. However, they did not measure the recall of the manipulated pictorial elements, but rather the recall of the information presented as

consistent in both conditions (Schüler, 2017, 2019). Therefore, it might be that discrepancies specifically affect the recall of the manipulated pictorial elements - either discrepant or not discrepant - whereas the recall of all other pictorial elements is not affected. This assumption is also supported by the finding of Experiment 1 that the paintings' elements are fixated longer when they are described as being discrepant to the historical evidence, compared to when they are described without naming the discrepancy.

Finally, Experiment 2 aimed to replicate the effect of named discrepancies on transportation with the transportation-short-scale (Appel et al., 2015), which includes six items with a corroborated factorial structure, reliability, and validity, instead of the single item question used in Experiment 1.

Taken together, we hypothesized that the paintings are perceived as being less trustworthy documents when discrepancies are named compared to when no discrepancies are named by the audio-text, but that this effect should be compensated by explaining the benevolent intentions of the artist (H1). We further hypothesized better retention of the manipulated pictorial elements when these elements are named as discrepant compared to when these elements are named without mentioning the discrepancy (H2). In addition, we exploratively analyzed the retention of pictorial elements that were named as consistent in all conditions. Based on prior research on narrative processing and the results of Experiment 1, we again hypothesized that transportation into the depicted historic narrative is lower with the discrepancy named compared to without the discrepancy named (H3).

Method

Participants

We recruited 133 participants from our institute’s mailing list. Fourteen were excluded because they were familiar with one or more paintings used for the analysis. Seven additional participants were excluded because they mixed up the paintings in the free recall test. Further, one participant was excluded because he or she did not speak German on a

native speaker level, and one was excluded due to technical issues. For the analysis, *n* = 110 participants remained (95 female, 14 males, 1 diverse; aged between 19 and 34 years, *M* = 23.96, *SD* = 3.14). They were from broad fields of studies (e.g. psychology, theology, medicine, media studies, natural sciences) and were randomly assigned to one of our four experimental conditions (*n* = 27 without discrepancy and without explanation, *n* = 28 with discrepancy and without explanation, *n* = 29 without discrepancy and with explanation, *n* = 26 with discrepancy and with explanation).

Design

We used a 2x2x3 design with the between-subjects factors discrepancy (with discrepancy named vs. without discrepancy named) and explanation (with explanation vs. without explanation), and the within-subjects factor painting (West, Hellqvist, Leutze).

Material

Compared to Experiment 1, we presented a sequence of nine instead of seven historical paintings on a computer screen. The first painting was “The Surrender of Breda” (1635) by Diego Velázquez, which is exhibited in the Museo del Prado in Madrid, followed by “The Founding of the Nation” (1967) by Dong Xiwen. After this, the three paintings used for the analysis followed: “Washington Crossing the Delaware” (1851) by Emanuel Leutze, “The Death of General Wolfe” (1770) by Benjamin West, and “Waldemar Atterdag Holding Visby to Ransom” (1882) by Carl Gustaf Hellqvist. The last four paintings of the presentation were “The Babylonian Marriage Market” (1875) by Edward Long, exhibited in the Royal Holloway College in London, “The Constitution of the 3rd of May 1791” (1891) by Jan Matejko which is exhibited in the Royal Castle of Warsaw, “The Proclamation of the German Emperor” (1885) by Anton von Werner, exhibited in the Otto-von-Bismarck-Stiftung in Friedrichsruh, and lastly, “The Death of Socrates” by Jacques-Louis David, exhibited in the Metropolitan Museum of Art in New York. The paintings used for the analysis (West, Hellqvist, Leutze) were chosen because they contained at least four pictorial elements that were consistent and four elements that were discrepant to a more plausible version of the historical event, as far as it

can be anticipated based on more reliable historic sources. These three paintings were also part of the test material in Experiment 1. Because of the high retention in Experiment 1, we presented the test paintings not at the end of the presentation but in the middle to make the free recall test more difficult.

Each painting was accompanied by an audio-text describing and interpreting the painting. For each of the three paintings used for the analysis, four different versions of the audio-text were created, depending on the condition. The audio-texts in all conditions named eight pictorial elements for each of the test paintings and therefore twice as many as in Experiment 1. Four of these pictorial elements were always named without discrepancy in all conditions (further referred to as the always consistent elements). The naming of the other four elements was manipulated according to the condition. The audio-text named these four manipulated pictorial elements either as discrepant (with discrepancy named) or not (without discrepancy named) and provided an explanation for the discrepancy (with explanation) or not (without explanation). The explanation was formulated in a way that the argumentation also made sense when no discrepancy was named. When no explanation was given, a text of similar length and with similar content was provided, which did not explain the discrepancy. Therefore, the audio-texts were equal in duration and only differed in a few words (see Table 4 for an example). For each of the paintings that was not used in the analysis, we also created four versions, depending on the conditions, to keep the naming or not naming of discrepancies with either explanation in each condition consistent. However, the audio-texts for the paintings that were not used for the analysis were shorter and named only one or two historical inaccuracies of each painting.

Table 4

Example of Text for the Native American in Wests Painting no Explanation Given vs. an Explanation Given after Discrepancy Named or not Named

Without explanation	With explanation
Benjamin West painted the picture 11 years after the event in London for an English audience. The appearance of the North American Iroquois was not very well known at that time and the viewers could therefore not recognize North America as the place of the action in Leutze's painting.	Benjamin West painted the picture 11 years after the event in London for an English audience. The appearance of the North American Iroquois was already very well known at that time and Leutze helped the viewers to recognize North America as the place of action with his depiction.

Note. Benevolent intention of the artist is underlined.

Technical equipment

The audio-visual presentation was done with the software IWM Study, which was installed on four computers with a 23-inch monitor [1920x1080px]. Participants navigated with the mouse through the experiment. The audio-texts were presented via headset.

Measures

Trustworthiness. Trustworthiness of each painting was measured with a translation of the five-item scale (Flanagin & Metzger, 2003) on a 7-point Likert scale, ranging from one (not at all) to seven (very much). One item for example was: “How credible do you think the painting is?” We calculated the mean score of trustworthiness for each painting. Cronbach’s Alpha for the trustworthiness rating of the Hellqvist painting was $\alpha = .89$, for the West painting $\alpha = .91$, and the Leutze painting $\alpha = .89$.

Retention. To measure retention, we used the same test as in Experiment 1, with a slightly shortened time of seven minutes for answering each test of the paintings. We calculated the

mean scores for the retention of the manipulated pictorial elements and the mean for the retention of the always consistent pictorial elements for each painting. The answers in the tests were rated by two independent raters with an interrater reliability (Krippendorff's Alpha) of α = .93 for the manipulated pictorial elements. The interrater reliability for the always consistent elements was α = .90. A third rater decided in cases when the two raters did not agree.

Transportation. The transportation into the events depicted in the paintings was measured with an adaptation of the six-item short-scale (Appel et al., 2015). The 7-point Likert scale ranged from one (not at all) to seven (very much). One item for example was "While viewing the painting I could imagine […] vividly". For […] we used the description of one of the manipulated pictorial elements as how it is depicted in the painting, for example "the rearing horses" in the Leutze painting. We calculated the mean score of transportation for each painting. Cronbach's Alpha for the transportation rating of the Hellqvist painting was α = .84, for the West painting α = .78, and the Leutze painting α = .80.

Procedure

Up to four participants were tested at the same time. Participants were welcomed and seated in front of the computer screen. They received the following instructions: "In the following you can view paintings depicting historical events. Each painting is accompanied by an audio-text providing further information about the painting. After the presentation of the paintings has ended, you will have to answer questions about these paintings". After the instructions, the nine paintings were presented sequentially with the respective audio-text. Each painting was preceded by a written introduction stating the title of the painting and the artist. The participants continued to the paintings by pressing the space bar. In contrast to Experiment 1, the audio-texts were no longer presented in self-paced segments but started and ended together with the presentation of the respective painting. This ensured that all participants had equal viewing times. After the presentation of each painting, the questions about transportation and trustworthiness were given. The three paintings used for the analysis were

presented in a balanced order for all possible sequences in all conditions to prevent order effects.

After the audio-visual presentation of all paintings had ended, the same filler task as in Experiment 1 followed again for seven minutes. Then the free recall tests for the three paintings were handed out successively. The participants had seven minutes for the free recall test of each painting. The testing sequence of the recall was balanced for the three paintings and matched the presentation sequence. Lastly, the participants answered the demographic questions and whether they had already known any of the paintings. At the end, they were debriefed and paid 10 Euros for their participation. The study received institutional research ethics committee approval.

Results

Trustworthiness of the Paintings

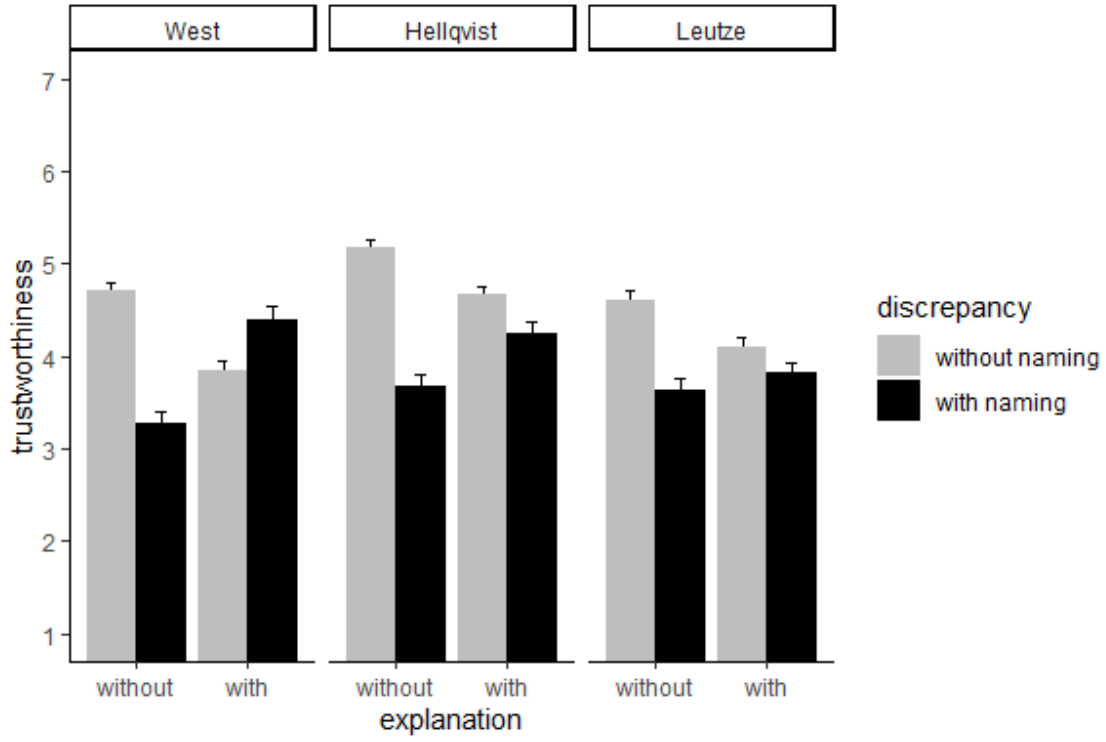
The ANOVA for trustworthiness revealed a significant main effect of discrepancy, $F(1, 106) = 13.14$, $p < .001$, $\eta^2_p = .110$, whereas the main effect of explanation was not significant, $F(1, 106) < 0.01$, $p > .999$, $\eta^2_p < .001$.. The main effect was qualified by a significant two-way interaction between discrepancy and explanation, $F(1, 106) = 11.07$, $p = .001$, $\eta^2_p = .095$. The trustworthiness rating was significantly lower with discrepancies named ($M = 3.53$, $SD = 1.15$) than with discrepancies not named ($M = 4.83$, $SD = 0.82$) if no explanation was given, $p < .001$. In contrast, the trustworthiness rating did not differ between discrepancies named ($M = 4.15$, $SD = 1.07$) and not named ($M = 4.27$, $SD = 0.84$) if an explanation was given, $p = .833$.

The three paintings also had an effect on perceived trustworthiness. There was a main effect of painting, $F(2, 212) = 8.32$, $p < .001$, $\eta^2_p = .073$, with trustworthiness of the Hellqvist painting being rated higher ($M = 4.44$, $SD = 1.23$) than the West painting ($M = 4.04$, $SD = 1.33$), $p = .002$, and the Leutze painting ($M = 4.04$, $SD = 1.24$), $p = .002$. While no significant two-way interactions between painting and discrepancy, $F(2, 212) = 2.77$, $p =$

.065, η^2_p = .025, and between explanation and painting, $F(2, 212) = 0.87$, p = .421, η^2_p = .008, were found, the three-way interaction between discrepancy, explanation and painting was significant, $F(2, 212) = 4.58$, p = .011, η^2_p = .041. For all paintings, Bonferroni-adjusted comparisons showed a similar pattern, namely, that the trustworthiness was lower in the condition with discrepancy named compared to the condition without discrepancies named if no explanation was given, whereas with an explanation, there was no significant difference between the condition with discrepancy named and the condition without discrepancies named However, giving an explanation when the discrepancy was named increased trustworthiness significantly for the West painting (p = .011), but only descriptively for the Hellqvist (p = .088), and the Leutze painting (p = .132). When the discrepancy was not named, giving an explanation lowered the trustworthiness significantly for the West painting (p = .001) but not for the Hellqvist (p = .073), and the Leutze painting (p = .585). In sum, the results show that the naming of discrepancies lowers the perceived trustworthiness of the paintings and that this effect is compensated by an explanation for the discrepancy.

Figure 1

Interaction Between Discrepancy and Explanation with Regard to Trustworthiness



Note. The error bars indicate standard errors.

Retention of the Manipulated Pictorial Elements

The ANOVA for the retention of the manipulated pictorial elements revealed neither a significant main effect of naming the discrepancy, $F(1, 106) = 0.37$, $p = .544$, $\eta^2_p = .003$, nor a significant main effect of explanation, $F(1, 106) = 0.32$, $p = .571$, $\eta^2_p = .003$, and also no significant two-way interaction between discrepancy and explanation, $F(1, 106) = 0.08$, $p = .776$, $\eta^2_p = .001$.

The main effect of painting was significant, $F(2, 212) = 65.34$, $p < .001$, $\eta^2_p = .381$. The retention for the manipulated pictorial elements was worse for the West ($M = 64.55$, $SD = 22.05$) than for the Leutze ($M = 86.59$, $SD = 20.23$), $p < .001$ and for the Hellqvist elements ($M = 89.09$, $SD = 14.97$), $p < .001$. The retention for Hellqvist and Leutze did not significantly differ, $p = .764$. Painting did not significantly interact with discrepancy and explanation, neither in the two-way interactions (discrepancy by painting, $F(2, 212) = 0.71$, $p = .494$, $\eta^2_p = .007$; explanation by painting, $F(2, 212) = 0.36$, $p = .698$, $\eta^2_p = .003$), nor in the three-way interaction between discrepancy, explanation, and painting, $F(2, 212) = 1.54$, $p = .218$, $\eta^2_p = .014$.

Retention of the Always Consistent Pictorial Elements

We also looked exploratively at the retention of those pictorial elements that were named as always consistent with the historical evidence in all conditions. The ANOVA revealed a significant main effect of discrepancy, $F(1, 106) = 5.85$, $p = .017$, $\eta^2_p = .052$. The recall of the always consistent pictorial elements was lower in the condition with discrepancies named ($M = 63.73$, $SD = 18.47$) compared to the condition without discrepancies named ($M = 72.02$, $SD = 17.58$). Both the main effect of explanation, $F(1, 106) = 0.44$, $p = .508$, $\eta^2_p = .004$, and the two-way interaction between discrepancy and explanation , $F(1, 106) = 0.15$, $p = .704$, $\eta^2_p = .001$, were not significant.

The main effect of painting was significant, $F(2, 212) = 38.23$, $p < .001$, $\eta^2_p = .265$. The always consistent pictorial elements of the West ($M = 73.18$, $SD = 24.35$) and the

Hellqvist painting (*M* = 75.00, *SD* = 22.46) were recalled better than those of the Leutze painting (*M* = 55.68, *SD* = 23.62). Painting did not significantly interact with discrepancy and explanation, neither in the two-way interactions (discrepancy by painting, $F(2, 212) = 0.52$, $p = .594$, $\eta^2_p = .005$; explanation and painting, $F(2, 212) = 0.01$, $p = .993$, $\eta^2_p < .001$.), nor in the three-way interaction between discrepancy, explanation and painting, $F(2, 212) = 0.31$, $p = .733$, $\eta^2_p = .003$. Thus, the results of the retention of the always consistent pictorial elements of all three paintings point in the same direction with lower retention with discrepancies named than without discrepancies named and showed that with discrepancies named, the retention of the always consistent elements was lower than without the naming of discrepancies.

Transportation

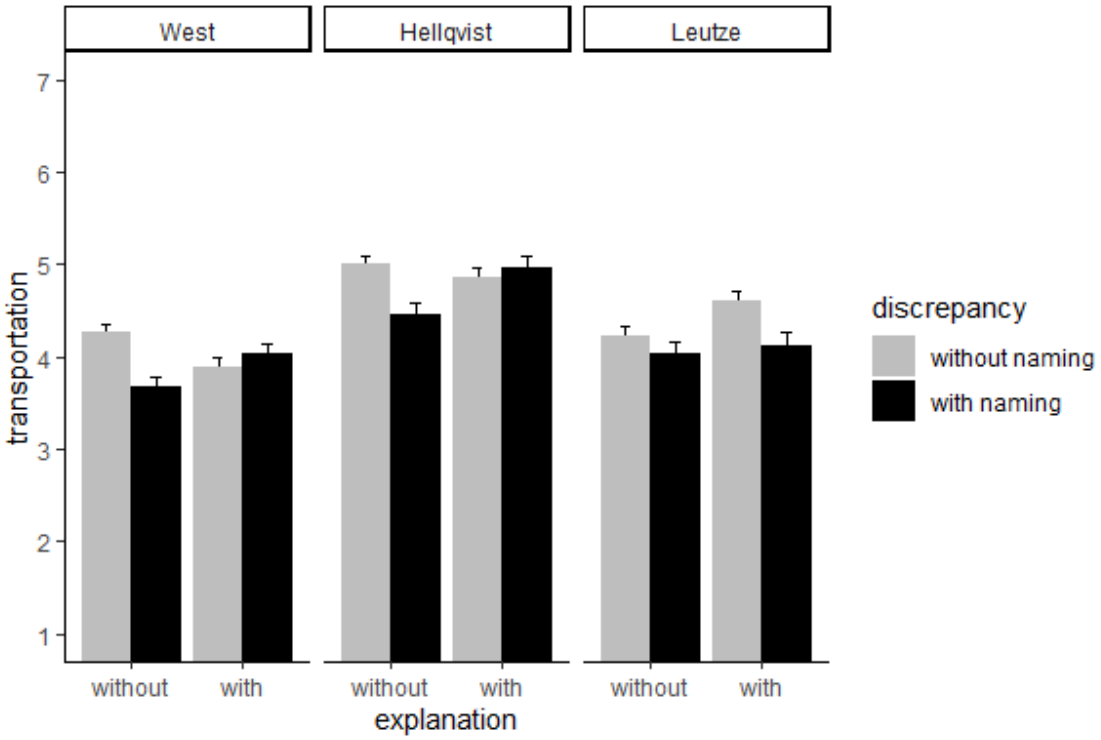
The ANOVA for transportation revealed no significant main effect of discrepancy, $F(1, 106) = 1.80$, $p = .183$, $\eta^2_p = .017$, no significant main effect of explanation, $F(1, 106) = 0.50$, $p = .481$, $\eta^2_p = .005$, and also no significant interaction between the two factors, $F(1, 106) = 0.94$, $p = .334$, $\eta^2_p = .009$.

There was a significant main effect of painting on transportation, $F(2 ,212) = 40.72$, $p < .001$, $\eta^2_p = .278$. The transportation for Hellqvist (*M* = 4.82, *SD* = 1.17) was higher than for Leutze (*M* = 4.26, *SD* = 1.20), $p < .001$ and West (*M* = 3.96, *SD* = 1.08), $p < .001$. The transportation for Leutze was significantly higher than for West, $p = .007$. While the two-way interactions between painting and discrepancy, $F(2, 212) = 0.27$, $p = .768$, $\eta^2_p = .002$, and between painting and explanation, $F(2, 212) = 0.83$, $p = .439$, $\eta^2_p = .008$, were not significant, a significant three-way interaction between discrepancy, explanation and painting was found, $F(2, 212) = 4.23$, $p = .016$, $\eta^2_p = .038$. Bonferroni-adjusted comparisons showed significant differences between conditions only for the West painting, whereas for the Hellqvist and the Leutze paintings no significant differences between the conditions were found. In the West painting, transportation was significantly lower in the condition with discrepancy named (*M* = 3.68, *SD* = 1.20) compared to the condition without discrepancies named (*M* = 4.26, *SD* =

0.92), if no explanation was given (p = .050), whereas with explanation there was no significant difference (p = .631) between the condition with discrepancy named (M = 4.03, SD = 1.17) and the condition without discrepancies named (M = 3.89, SD = 1.03).

Figure 2

Interaction Between Discrepancy and Explanation with Regard to Transportation



Note. The error bars indicate standard errors.

Discussion

Experiment 2 investigated how the explicit naming of the historical inaccuracies in a historical painting and the provision of either an explanation or no provision of an explanation for these inaccuracies influences the trustworthiness, the retention, and the transportation into the depicted narrative of a historical painting. The results support H1; that is, the trustworthiness is rated lower when discrepancies are named, and this effect is compensated by explaining the discrepancy with the benevolent intention of the artist. The results do not support Hypothesis H2, meaning that the retention of the manipulated pictorial elements is higher with discrepancies named compared to without discrepancies named. But, surprisingly, the explorative analysis showed worse retention of the always consistent

elements in the condition with discrepancies named compared to without. We also did find weak support for Hypothesis H3, as transportation was lower with discrepancies named compared to without, but only in one of three paintings and only if no benevolent explanation was available.

The compensating effect of the benevolent explanation on perceived trustworthiness indicates that viewers do indeed consider information about the artist when it is helpful to explain a conflict. This effect is in line with the literature on multiple conflicting text documents (Gottschling et al., 2019; Kammerer et al., 2016; Saux et al., 2018). Our study extends this sourcing effect to situations with explicitly named discrepancies between text and picture documents.

In contrast, the results of the free recall task did not correspond to the initial hypotheses. While it was expected that discrepancies would affect the recall of the manipulated pictorial elements, either discrepant or not discrepant, but not the recall of all other pictorial elements, the reverse pattern was observed. First, the finding that pictorial elements were not better retained if they were named as discrepant to the historical evidence differs from the results of previous text studies (Albrecht & O'Brien, 1993; O'Brien & Myers, 1985) and is also at odds with the findings of Experiment 1, which showed longer processing of the elements that were named as discrepant to the historical evidence. Second, the finding that the always consistent pictorial elements were better recalled in the condition in which other elements were named as discrepant is in line with the first experiment by Schüler (2017) but not compatible with Experiment 2 by Schüler (2017) and also not with the results in Schüler (2019). One tentative explanation could be that cognitive resources were shifted away from memorizing consistent content towards the processing of discrepant content, thereby hampering the construction of an encompassing mental model of the painting's content. But it should be kept in mind that the analysis of the recall of the always consistent pictorial elements was explorative in nature and should not be overinterpreted.

Finally, an effect of naming or explaining discrepancies between a painting and the historical evidence on perceived transportation was found only in one of three paintings. Hence, we could not fully replicate the effect of discrepancy on transportation in Experiment 1 by using the transportation short scale. A shift away from a narrative processing due to the naming of discrepancies between a painting and a more plausible version of the historic event is therefore only partly supported.

General Discussion

Pictures as well as texts constitute important historical sources (Burke, 2001). However, despite the seminal study of Wineburg (1991), research has mainly focused on the processing of texts, whereas only a few studies have investigated historical reasoning using picture documents (Smith et al., 2019; van Boxtel & van Drie, 2012). Therefore, our study investigated multiple documents of mixed media formats, namely, paintings of historical events together with audio-texts. We were interested in how discrepancies between the depiction of the historic event by the paintings and today's most plausible versions of the historic events commented on in the audio texts will affect the recipients' cognitive processing. The audio-texts systematically varied whether or not historical inaccuracies of the paintings were explicitly named (Experiments 1 and 2) and whether or not the inaccuracies were explained by pointing out the benevolent intention of the artist (Experiment 2). Based on the literature, we expected conflict recognition indicated by attention to the pictorial elements named as inaccurate together with different conflict regulation strategies indicated by the attention on the source label or by the interest in further information in Experiment 1, the trustworthiness of the paintings in Experiment 2, as well as the retention of pictorial elements in Experiments 1 and 2. We further expected that discrepancies lower the experience of transportation in both experiments.

The results support the assumption that confronting viewers of historical paintings with discrepancies between the paintings and the most plausible version of the historical event triggers attempts to restore coherence. In Experiment 1, closer inspection of discrepant

pictorial elements was indicated by increased fixation times on pictorial elements if their discrepancy to the historical evidence was named. This is in line with research on discrepancies in single texts (e.g., Albrecht & O'Brien, 1993), multiple texts (Beker et al., 2016), and between text and picture (Schüler, 2017, 2019) reporting longer processing times of conflicting compared to consistent information. Although we found longer processing of the picture elements that were named as discrepant in Experiment 1, we did not find a corresponding memory advantage for these elements in Experiment 2. In addition, the explorative analysis of Experiment 2 revealed that the memory for the always consistent picture elements was lower due to the naming of some pictorial elements as discrepant. This could indicate that cognitive resources were shifted away from memorizing consistent content towards an intensified processing of the discrepant content. This pattern of results is partly compatible with previous studies. The studies using text material observed positive effects of discrepancies on the retention of the discrepant information (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995). In two experiments using text and picture material, there was no difference in the retention of consistent pictorial elements with or without a discrepancy present (Experiment 2 by Schüler, 2017; Schüler 2019), but in one experiment, a worse retention of the consistent pictorial elements similar to our findings was reported (Experiment 1 of Schüler, 2017).

We also found mixed evidence for the conflict regulating strategy of taking background information about the document's source into account. Viewers did not pay more attention to the source label if a painting's discrepancy with the historical evidence was named in Experiment 1, although longer processing of the information about the author due to conflicting information has been frequently shown in multiple text research (Braasch et al., 2012; Bråten & Braasch, 2018). One possible explanation for this lack of attention might be that most viewers read the label at the beginning of the presentation and still remembered the information when the discrepancy was named. It may also be the case that the audio-texts were perceived as superior commentators on the historical accuracy of the paintings. This may have led to a predefined hierarchy of the multiple documents with regard to their

quality as a historical source and, therefore, there was no need to evaluate the sources of the paintings. For this reason, we presented either a benevolent explanation or no explanation and examined trustworthiness ratings as an alternative indicator of sourcing in Experiment 2. In this way, we could observe that viewers rated the trustworthiness lower for the painting when discrepancies were named than when they were not named and that this effect was compensated by a benevolent explanation that was useful to explain the discrepancy. This result supports the assumption that viewers encountering discrepancies use the information about the authors to regulate the conflict when it is useful to do so. Our result regarding sourcing is consistent with the literature using text material (Gottschling et al., 2019; Saux et al., 2018). Furthermore, our result extends this effect to situations in which one source is commenting on the other source and a discrepancy is explicitly named.

Concerning the third possible conflict regulation strategy, we did not find a higher interest in further information about the painting with discrepancies named compared to without discrepancies named, indicating that the viewers did not postpone the conflict regulation. In the light of the models of question asking, stating that conflicting content leads to more question asking than content without conflicts, our result regarding interest is unexpected (Graesser & McMahen, 1993; Otero & Graesser, 2001). On the one hand, it might be that the participants were able to regulate the conflict after encountering the discrepancies in such a way that they were satisfied with the resolution and therefore developed no increased interest in further information. On the other hand, it might also be that the viewers were not fully satisfied with the results of their conflict regulation, but instead of becoming interested in further information, they became confused about the facts of the historic event. Interest arises from novel information that can be understood by the individual, while confusion arises from novel information that cannot be easily understood (Silvia, 2013). Future research could investigate more deeply the emotional experience linked to the naming of discrepancies or could measure question-asking as an indicator of postponing the resolution of conflicts.

Together with the finding that more than one half of the participants mentioned the conflict in the free recall task of Experiment 1 without being asked, the pattern of results indicates that the participants became aware of the discrepancies and engaged in strategies of conflict regulation, particularly by considering background information about the artist's intentions.

Lastly, regarding transportation, we observed that discrepancies reduced the viewer's experienced transportation in Experiment 1. However, we could not fully replicate this finding in Experiment 2 using a more elaborated measure of transportation. In Experiment 2, the transportation was significantly lower with discrepancies named compared to without just for one of three paintings and only when no explanation was given. For the other two paintings, the observed effect was in the same direction but did not reach a significant level. Besides using a short scale instead of a single item, the formulation of the transportation task also differed slightly between the two experiments. Whereas the participants in Experiment 1were asked about transportation into the historic event, in Experiment 2, the participants were asked about transportation into the event depicted in the painting. For example, to measure transportation for the Hellqvist painting, we asked how vividly they could imagine a married woman of Visby wearing her hair uncovered as it was depicted instead of asking about their vivid imagination of married women wearing their hair covered as was probably the case in medieval Visby. It can be argued that these questions led to different judgements since the historic event needs the imagination of the participants, whereas the depicted event is visible to the participants. Hence, the painting's inaccuracies are likely to be more relevant for hampering the transportation into the historic event than the depicted event itself. This difference in the measurement might have narrowed the effect on transportation in Experiment 2. Nevertheless, the study provides only weak evidence for an effect of discrepancies on the experienced transportation as proposed by the model of narrative engagement (Busselle & Bilandzic, 2008), at least for static depictions of historical events.

Taken together, the results of our study using pictures and audio text documents support assumptions of the CSI model with media types different from texts. Generalization of findings from written text to other media types is not trivial, since attributes of media known from multimedia research might influence the three stages proposed by the CSI model, namely conflict detection, regulation, and resolution. At least five attributes can thereby play a role: Pictures in contrast to texts are not processed linearly from beginning to end and some pictorial elements or conflicting elements can be overlooked more easily if they are not directly addressed. Pictures in contrast to text have a higher potential for immersion of the recipient, which could also affect the detection of conflicts according to the model of narrative processing (Busselle & Bilandzic, 2008). Audio information is more transient compared to written text (Wong et al., 2012), which could affect conflict detection and the memory for source information presented therein. However, the simultaneous processing of picture and audio text could enhance co-activation of conflicting information and therefore conflict detection compared to a sequential processing of written texts. Lastly, the modality might play a role, and especially text and picture combinations or video presentation have the potential for dual coding (Paivio, 1990). The document combination of picture and audio-text used in our experiment was very different from only text on all these attributes. Nevertheless, Experiment 1 showed that recipients detected the conflicts presented in picture and audio-text. In addition, Experiment 2 revealed that, although the relevant information about the source was presented in a transient audio-text, the recipients were able to use this information in order to evaluate the trustworthiness of the documents. It is also reasonable that the lower memory for consistent pictorial elements is a side effect of the transient presentation of information in the audio-text. It might be that conflict regulation captured cognitive resources that were subsequently not available for memorizing consistent pictorial elements presented in the next part of the audio-text. Therefore, this result might not be replicated using written texts, as texts can be processed in one’s own pace, and readers can pause their processing to think about a conflict without missing incoming information. Considering our results together with previous studies (List, 2018; List & Ballenger, 2019;

Merkt & Huff, 2020; Salmerón et al., 2020), investigating different media types and incorporating multimedia principles can be fruitful for the generalization and identification of boundary conditions of the framework of multiple documents and the CSI model.

Regarding the generalizability of the findings, some additional limitations must be noted. First, the present study focused on document pairs, with one document (the audio-text) explicitly referring to and commenting on the second one (the historical painting). While this is a rather common type of relation between documents, with examples ranging from letters to the editor in newspapers about certain articles to buyers' comments on books in online shops to explanations of documents in history textbooks, previous research on conflicting information and multiple documents has focused on other types of text sets. More specifically, text studies have often used incoherent behavior of a character (Albrecht & O'Brien, 1993), while multiple document research has often used conflicts, such as authors describing different causes for climate change without explicit cross-referencing (Strømsø et al., 2010). From a theoretical point of view, it would be interesting to investigate conflict regulation strategies with those more traditional types of document pairs, for example, a painting of a historic event together with an eyewitness report for which the viewers have to detect the inconsistencies on their own. Experiments closer to the original paradigm would provide additional evidence for those theories from text and multiple text document research that can be applied to picture documents. Second, in our experiments, we used historical paintings that can be used as historical sources but are also considered to be art. It would be interesting to know whether the compensating effect of a benevolent explanation would generalize to other material. Future research should therefore address whether perceiving an author as highly benevolent would also restore the trustworthiness for distorted pictures presented in the news, social media, advertisement, or even as propaganda. Third, we did not include a baseline measurement for interest in the presented historical topics or history or art in general in the two conditions. Hence, the participants may have differed with regard to their interest in information about the historical topic already at the beginning of the experiment.

Despite these limitations, the present study substantially extends previous research about conflict regulation with multiple documents from text sources to combinations of pictures, source labels, and audio-texts. We could show that viewers recognized conflicts by allocating additional attention to the conflicting information which is in line with evidence from text (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995), text-picture (Schüler, 2017, 2019), and multiple text literature (Beker et al., 2016). Furthermore, viewers use information about the artist to regulate explicit conflicts between text and picture documents, as they do use information about the author with conflicts between multiple text documents (Gottschling et al., 2019; Saux et al., 2018). This supports the conclusion that conflict processing and related conflict regulation strategies generalize from multiple text documents to text-picture documents and could apply also to documents of other (mixed) media formats. Besides its theoretical relevance, understanding these processes can be especially beneficial for the design of informal learning settings in museums of art and (art-) history as well as for the design of formal learning settings including history textbooks. Designers should consider that the naming of historical inaccuracies changes the gaze behavior of the viewers and that explanations can restore the perceived trustworthiness of a painting as a historical document after the painting's historical inaccuracies have been named.

Declaration according to § 5 Abs. 2 No. 8 of the PhD regulations of the Faculty of Science

-Collaborative Publications-

The following chapter (Chapter 3) consists of a pre-print version of a manuscript that is published. It was co-authored by Manuela Glaser and Stephan Schwan. The proportional contributions to this manuscript are presented in the subsequent table.

Author	Author position	Scientific ideas (%)	Data generation (%)	Analysis and Interpretation (%)	Paper writing (%)
Manuel Knoos	First Author	80	100	80	80
Manuela Glaser	Second Author	10	0	10	10
Stephan Schwan	Third Author	10	0	10	10
Title of the paper	Aesthetic Experience of Representational Art: Liking is affected by Audio-Information Naming and Explaining Inaccuracies of Historical Paintings				
Status in publication process:	Published. Knoos, M., Glaser, M., & Schwan, S. (2021). Aesthetic Experience of Representational Art: Liking Is Affected by Audio-Information Naming and Explaining Inaccuracies of Historical Paintings. *Frontiers in Psychology*, 3009. https://doi.org/10.3389/fpsyg.2021.613391				

3. Study 3

Aesthetic Experience of Representational Art: Liking is affected by Audio-Information Naming and Explaining Inaccuracies of Historical Paintings

Artworks in museums are often presented together with additional information, such as titles, text labels, or oral explanations in the form of personal or audio guides. In the past years a number of studies focused on the effects of titles on the aesthetic experience of paintings but did not examine the effects of longer additional information such as accompanying audio-texts. However, since art- and art-history museums do not change the titles but frequently provide their visitors with audio guides that include longer explanations of the paintings, examining the viewer's aesthetic experience of paintings in combination with longer accompanying audio-texts is both of theoretical and practical relevance. Such audio-texts are intended to educate the viewers, to help them to understand the artworks and thereby enhance the visitors' aesthetic experience in the gallery. They thus differ from other additional information, such as information about the prices of paintings or opinions of other people that could also influence the aesthetic experience, but in a different way, namely via priming the viewers expectations (Lauring et al., 2016).Research corroborates positive effects of titles and short text labels on the viewer's subjective understanding of paintings (Bubić et al., 2017; Leder et al., 2006; Russell, 2003; Swami, 2013) and their aesthetic appreciation (Belke et al., 2010; Bubić et al., 2017; Gerger & Leder, 2015; Millis, 2001; Russell, 2003; Swami, 2013). However, a recent review (Chmiel & Schubert, 2019) points out a substantial number of studies that did not observe effects of additional information on subjective understanding and aesthetic appreciation in the form of liking. Other research even shows negative effects of mismatching titles on liking (Belke et al., 2010; Gerger & Leder, 2015). For this reason, conditions need to be specified when and how additional information related to an artworks' meaning influences subjective understanding and liking of artwork as two main aspects of the aesthetic experience (Leder et al., 2004).

First, the effects of titles on subjective understanding and liking mostly apply for abstract rather than for representational art (Chmiel & Schubert, 2019). This has been corroborated by several studies showing effects for highly abstract but not for representational art (Leder et al., 2006; Moore & West, 2012; Swami, 2013). This could indicate, that the iconicity of representational art could provide the viewers with a feeling of an easy and high understanding, whereas abstract art needs clarification of what the painting represents. The absence of effects of additional information on the subjective understanding and liking of representational art might thus be due to the viewers feeling of an already highly subjective understanding and liking even if no additional information is provided.

Second, effects of additional information on the subjective understanding and liking of artwork may depend on the type of the additional information. Comparing descriptive and elaborative titles to a control group without titles, an experiment (Leder et al., 2006) revealed that both titles improved the subjective understanding of paintings compared to the control group. Elaborative titles had the highest effect on subjective understanding but neither of the titles increased liking. Comparing title, broad genre information, and content specific information to a control group without additional information, an experiment (Experiment 1 of Swami, 2013) found that all three types of information improved the subjective understanding of abstract paintings compared to the control condition, but content specific information had the highest effect and was the only type of information that improved liking. In addition, the type of additional information can influence whether the additional information affects the liking of a painting positively or negatively. Studies show that paintings are liked more when the provided title semantically matches the content of the painting than if the title does not semantically match the content and this mismatch remains unexplained (Belke et al., 2010; Gerger & Leder, 2015). One of the studies showed this for representational paintings (Belke et al., 2010) but when comparing the group with matching titles and the group with unexplained non-matching titles to the control condition that did not receive any titles, it can be concluded that this effect was mainly driven by representational paintings being less liked due to the unexplained mismatch of title and content than paintings being more liked due to a

match between the title and content. In other words, while the high liking typical for representational art is not easily enhanced by titles, unexplained inconsistencies such as mismatching titles can substantially decrease the liking of representational art. The authors (Belke et al., 2010; Gerger & Leder, 2015) assume that the reduced liking is caused either by lower processing fluency and meaning making or by a reduced understanding of the painting, higher liking is assumed to be caused by better understanding, higher processing fluency and disfluency reduction.

That unexplained inconsistencies can lead to a disfluent processing is supported by research outside the field of aesthetics. For example, discrepancies between a map and a related text led to longer fixation times on the text and the picture than text and map providing similar information, which was interpreted as a hampered process of information integration (Schüler, 2017). Disfluency due to unexplained inconsistencies might not only arise when additional information does not match the content of a picture or painting but also when the content of a seemingly realistic representational painting does not match reality. Historical paintings frequently contain historical inaccuracies, which are inconsistencies between the depiction of a historic event in a painting and a more plausible version of the event based on today's historians' opinions (Burke, 2001). Museums of art and history often provide additional information in the form of audio-texts naming the paintings inaccuracies. Without high background knowledge, the inaccuracies cannot be seen or inferred by looking at the painting. Therefore, mentioning them in the form of additional information is important for interpreting and understanding the painting. However, this could affect art processing and the evaluation of the representational artwork especially when the inconsistencies remain unexplained.

Art Processing, Aesthetic Emotions, and the Evaluation of Artworks

The studies on the subjective understanding and liking of art described in the previous section are mostly discussed in the context of two frameworks. First, the fluency theory (Reber et al., 2004) proposes that the easier the viewer's processing, meaning

making, and understanding of an artwork is, the more the artwork will be liked by the viewer. Fluency can thereby result from early processing stages such as the classification of the artwork or the perceptual analysis of symmetry but also from later higher order processing stages, such as the cognitive mastery of an artwork. Similar to the hypothesis of the fluency theory is the simplified hypothesis derived from the psycho-historical framework for the science of art appreciation (Bullot & Reber, 2013), stating that higher understanding of an artwork is positively linked to its aesthetic liking. In both frameworks, additional information, such as titles or explanations of the style and the art historical context, is assumed to enhance but also lower the liking of artworks, depending on whether the additional information contributes to a fluent processing and better understanding or to a less fluent processing and lower understanding.

The two frameworks have been recently incorporated into a more complex theoretical model proposing that positive and negative effects of additional information on the subjective understanding and liking of artworks as well as aesthetic emotions mainly depend on two cognitive appraisals made during the higher order processing stage of cognitive mastery of an artwork: The Vienna integrated model of art perception (VIMAP; Pelowski, Markey, Forster, Gerger, & Leder, 2017) proposes seven stages of art processing. The first stage is the pre-classification, which includes factors of context (museum, laboratory, social or individual setting) and personal factors (mood, personality, expectations) that influence the viewers processing and emerge before a person deals with an artwork. In the second stage, the perceptual analysis, the low-level features of an artwork are processed, such as complexity, contrasts, and color. In the third stage, the implicit memory integration, elements of the painting are combined to more or less meaningful patterns. Thereby, factors such as familiarity and prototypicality play a role. In the fourth stage, the explicit classification, viewers identify the content in accordance with the painting’s context, style, and information learned about the artist. In all these stages, the focus is mainly on bottom-up processes that influence the art perception of a viewer.

For effects of additional information on subjective understanding and liking, especially the fifth stage, the cognitive mastery, is important. Cognitive mastery is characterized by top-down processes that consider and combine the information gathered by the bottom-up processing in order to form coherent meaning of the artwork together with an appropriate evaluation and physical response. The outcome of this mastery process depends on two processing checks: schema congruency check and self-relevancy check. For the schema congruency check, viewers consider their schemas about their knowledge, expectations, understanding, and opportunities for learning (Silvia, 2009). Thereby they also consider the success of the processing during the former stages of basic perceptual processing, object identification, explicit classification, and integrating these elements. The match for each of these elements can be more or less congruent. A good overall match results in a subjective feeling of fluency and an efficient processing and understanding. For example, viewers could check whether their understanding of the artwork matches the level of understanding they expected. The second check proposed by the model is the self-relevancy check. With this check, the viewers consider the personal importance of the artwork for their self-image. The viewers decide whether the outcome of their viewing is relevant to them and whether they really have an interest or need to process the artwork. This is mainly the case for experts in real art situations but not for laypersons as the outcome of the layperson's art-processing does not threaten their self-image. Therefore, when considering a layperson's art processing, the self-relevancy check can be neglected, and the model then suggests two different outcomes based on the congruency-check.

The first outcome results from high schema congruency together with low self-relevance. It is characterized by a default or facile reaction. This is probably the most common outcome of viewers not finding something new or questioning in the artwork. The result is a sufficient classification, easy processing and understanding of the artwork with little emotional engagement, and a facile feeling of pleasure.

The second outcome results from low schema congruency together with low self-relevance. It is characterized by a reaction of novelty and small insight due to a small incongruency in the congruency-check. Certain aesthetic emotions are thereby triggered, depending on whether or not the viewers are able to resolve the incongruency. Viewers can resolve the incongruency either by continuing their processing to find more information that contributes to a higher match, or they render the incongruency as irrelevant, or viewers can modify their schema by generalizing definitions, classes, or expectations to include the novel elements. Furthermore, viewers can accept the incongruency as a mystery and accept the ambiguity and not seek a resolution, or the incongruency is explained by further additional information. In these cases, the viewers might find pleasure and interest for the incongruency. Alternatively, the viewers appraise the chance of finding a resolution to be low and experience a need for a resolution in order to restore coherence. This will result in confusion, in lower interest, and probably in a lower subjective understanding and liking of the artwork. Hence, confusion and interest are opposite outcomes of the same cause, depending on the viewer's appraised chance to form a coherent understanding after an incongruency was encountered (cf. Silvia, 2009). Independent of the appraised chance to solve the incongruency, all viewers should experience surprise when confronted with incongruency.

Empirical evidence of effects of additional information on aesthetic emotions is scarce. The present literature has mostly reported no effects of additional information in the form of titles on the emotional experience (Bubić et al., 2017) and interest (Gerger & Leder, 2015; Leder et al., 2006). One study considered art appreciation as a scale of interest and liking ratings together and reported significant effects of additional information (Swami, 2013). To the best of our knowledge, no studies have investigated the effects of additional information on surprise and confusion. Therefore, we found there was a need to empirically test the assumptions of the VIMAP with regard to emotional outcomes. We thereby expected not only that an incongruency in the congruency-check can arise from inconsistencies, such

as titles semantically mismatching the content, but also by naming the historical inconsistencies of the content of a representational painting mismatching reality.

Transportation as a Result of Processing a Narrative Artwork

As we used historical paintings in our study that depict a story and are therefore inherently narrative, we also considered theories of narrative processing to investigate the influence of additional inconsistent information on processing outcomes. According to the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008), the fluent processing of a narrative, presented either in the form of texts, films, or pictures, leads to the feeling of being transported into the story. The phenomenon of being transported is described as the readers' or viewers' experience of being mentally absorbed in the story world (Gerrig, 1993) and consists of a cognitive, emotional, and imaginary component (Green and Brock, 2002). Transported individuals focus their cognitive processing on the events of the story; they identify and feel with the characters and create vivid mental images of the places and characters. They can experience a flow-like state and lose awareness of what is going on around them (Green & Brock, 2000). Transportation is enjoyed by the recipients (Bilandzic & Busselle, 2011; Busselle & Bilandzic, 2008) and is therefore an essential experience also when processing narrative artworks. Transportation was mostly investigated with written text and movies but is assumed to apply to narratives presented in all modalities (Green & Brock, 2000). The model of narrative engagement assumes that readers and viewers fluently process and experience transportation when the story is coherent. However, when the recipients encounter incoherence or implausibilities that are not explained by the story world, processing fluency is diminished and transportation is lowered (Busselle & Bilandzic, 2008). This link between perceived realism and transportation is supported by empirical results (Bilandzic & Busselle, 2011; Green, 2004). Based on the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008) and the related empirical studies, we expected that naming a paintings’ inconsistencies reduces

transportation and that explaining these inconsistencies by benevolent intentions of the painter will compensate for this negative effect.

The Present Study

In our study, we investigated the effects of additional information naming a painting's inconsistencies on the viewers' art evaluation and aesthetic emotions when viewing representational art, that is, in specific historical paintings. Based on the VIMAP and related empirical findings (Belke et al., 2010; Gerger & Leder, 2015), we expected that the naming of a painting's inconsistencies and leaving them unexplained lowers the subjective understanding and the liking of the historical painting. Additionally, we assumed that informing the viewers about the artists' intentions in order to explain these inconsistencies can help the viewer to restore coherence. Hence, the information about the artists' intentions should compensate for the negative effects of naming inconsistencies without explaining them. This should be manifest in a two-way interaction between the factors naming of inconsistencies and explanation for subjective understanding (P1) and liking (P2) of the historical paintings. Based on the VIMAP, we further expected that surprise will generally be higher with inconsistencies named compared to without inconsistencies named, indicated by a main effect of naming inconsistencies (P3). Additionally, the viewers should experience lower interest and higher confusion with inconsistencies named compared to without inconsistencies named when no explanation is given, but these effects on interest (P4) and confusion (P5) should be compensated for by the provision of an explanation about the artists' intentions. Lastly, based on the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008), we expected that naming the painting's inconsistencies without explaining them reduces the viewers experienced transportation compared to not naming and explaining the inconsistencies, but explaining these inconsistencies by mentioning the artists' intentions should compensate for this effect (P6). Hence, for Predictions 4, 5, and 6 we again predicted two-way interactions between the factors naming of inconsistencies and explanation.

Method

Participants

The experiment was done online and could be accessed with all common browsers. We recruited 196 participants on Prolific and instructed them only to participate via computer or tablet and not via smartphone due to the small screen size, which we considered insufficient for noticing the details of the paintings and for an appropriate aesthetic experience of the paintings. The available participants were pre-filtered to include only native speakers of German. From the 196 participants, 41 participants were excluded because they already knew at least one of the three paintings we used in the present study. Four were excluded because they participated via smartphone. Six were excluded because they gave 50% or less correct answers in a memory check, indicating that they had guessed the answers and had not listened closely to the audiotexts commenting on the picture. The memory check presented a statement about the historic event, for example, that it was summer when Washington crossed the Delaware and asked if this was depicted in the painting or not. Four participants were excluded because they studied or worked in the field of history or art-history. Subsequently, 139 participants remained for the analysis: 64 (46 %) females, 75 (54 %) males; aged between 18 and 67 years (M = 30.78, SD = 9.95).

Design

We tested our predictions using a 2x2 design with naming of inconsistencies (with vs. without) and explanation (with vs. without) as the between-subjects factors. The 139 participants were randomly assigned to one of our four conditions (I-E-: without inconsistencies named and without explanation, n = 40; I+E-: with inconsistencies named and without explanation, n = 30; I-E+: without inconsistencies named and with explanation, n = 40; I+E+: with inconsistencies named and with explanation, n = 29).

Material

As research material, we used pictures of three historical paintings: “Valdemar Atterdag Holding Visby to Ransom” by Carl Gustaf Hellqvist, “The Death of General Wolfe” by Benjamin West, and “Washington Crossing the Delaware” by Emanuel Leutze. All of these paintings contain pictorial elements that are consistent and pictorial elements that are inconsistent with a plausible version of the historical event based on today’s historians’ opinions.

We created four different versions of audio-texts for each painting, depending on the respective condition. The audio-texts in all conditions commented on eight pictorial elements for each painting. This consisted of information about the location of the pictorial element in the painting, its description, and an interpretation of the element regarding the historic event. The audio-text’s interpretation of four pictorial elements of each painting was manipulated according to the condition. These elements were either named as being inconsistent to the actual historic event (with inconsistencies named) or not (without inconsistencies named). Directly after this, either information about the intention of the artist followed that was able to explain the inconsistency (with explanation) or a text of similar length and verbal content followed that did not inform the participants about the intention of the artist and did not explain the inconsistency (without explanation). The information about the artists' intentions was formulated in a way that made sense even when no inconsistencies were named. The intention of the artist was always benevolent, for example, by stating that the artist wanted to make a certain point clearer to the viewer (see Table 5 for an audio-text example). The audio texts had different durations for the paintings (4:29 min for Hellqvist, 4:41 min for Leutze, 4:35 min for West) but were of equal length for the four conditions with only minor changes in the sentences.

Table 5

Example of Text for the Native American in Wests Painting "The Death of General Wolfe" for the Four Conditions. Either the Inconsistency was Named or Not, then Either an Explanation Followed or Not.

On the left kneels a Native American wearing loincloth and a red feather. It is one of the Iroquois who were allied to the British. The Iroquois were engaged as scouts before the battle	
Without inconsistencies named	*With inconsistencies named*
During the combat they did indeed leave the camp and took part in the battle.	During the combat they did not leave the camp and did not take part in the battle.
Without explanation	*With explanation*
Benjamin West painted the picture in London 11 years after the event for an English audience. The appearance of the North American Iroquois was not very well known at that time, and the viewers could therefore not recognize North America as the place of the action in Leutze's painting.	Benjamin West painted the picture in London 11 years after the event for an English audience. The appearance of the North American Iroquois was already very well known at that time and Leutze helped the viewers to recognize North America as the place of action with his depiction.

Note. Intention of the artist explaining the discrepancy is underlined.

Measures

To control for a priori differences between the conditions, we measured the participants' general interest in art using the respective part of the German version of the Vienna Art Interest and Art Knowledge Questionnaire (Specker et al., 2018). Participants answered the questions on a 7-point Likert scale ranging from one (not at all) to seven

(completely) for their self-reported interest and from one (less than once a year) to seven (once a week or more) for their self-reported activities in the context of art. We calculated the mean score for general interest in art. The internal consistency of the general interest in art scale was good as indicated by a Cronbach's Alpha of α = .88.

We measured the subjective understanding for each painting with a two-item scale. Answers had to be given on a 7-point Likert scale ranging from one (not at all) to seven (very much). These items were similarly used by Swami (2013) and adapted from Silvia (2005). We calculated the mean score of subjective understanding. The internal consistency of the subjective understanding scale was good, as indicated by a Cronbach's Alpha of α = .88.

We measured liking, surprise, interest, and confusion for each of the three paintings with the two items of the respective sub-scales of the German Version of the Aesthemos scale (Schindler et al., 2017). We used the original instruction of the Aesthemos to focus the participants on their own aesthetic experience. The instruction states in German: Welche gefühlsmäßige Wirkung hatte x auf Sie? Bitte kreuzen Sie zu jedem Gefühl unten die Kategorie an, die auf Ihr persönliches Erleben am besten zutrifft. Bitte geben Sie nur an, wie Sie sich tatsächlich gefühlt haben. Beschreiben Sie nicht die Gefühle, welche im zuletzt gesehenen Gemälde ausgedrückt wurden, wenn Sie diese nicht selbst empfunden haben. [Which emotional effect did x have on you? For each emotion listed below, please mark the response category that best matches your personal experience. Please only indicate how you actually felt. Do not characterize the emotions expressed in x if you did not feel them yourself.] We replaced x with "das zuvor gesehene Gemälde [the previously seen painting]. For each emotion (liking, interest, confusion, surprise), answers were given on two items on a 5-point Likert scale ranging from one (not at all) to seven (very much). The two items of liking were "Empfand ich als schön [I found it beautiful]" and "Gefiel mir [I liked it]". We calculated the mean scores of liking, interest, confusion, and surprise. The internal consistency of all subscales of the Aesthemos were acceptable to good as indicated by

Cronbach's Alphas of α = .85 for liking, α = .88 for interest, α = .73 for confusion, and α = .80 for surprise.

We measured transportation into the historic event with the adapted version of the six-item transportation short-scale (Appel et al., 2015) after each painting. For item five and six stating "While viewing the painting, I could imagine […] vividly", we inserted one of the manipulated pictorial elements into the gaps, for example, "the behavior of the horses" for the Leutze painting. We calculated the mean score of transportation. The internal consistency of the transportation scale was excellent as indicated by a Cronbach's Alpha of α = .92.

Procedure

At the beginning of the experiment, the participants were instructed to focus on the paintings as artworks. For this, they were informed that they will see three paintings, of which the originals are exhibited in museums. Therefore, they can imagine the study to be similar to a visit in an art museum. The participants were further informed that the paintings depict historic events and that an accompanying audio-text will present further information about the painting, the artist, and the historic event. They were instructed that the audio-text for each painting could only be listened to once, and after viewing the paintings, they will be asked questions about the paintings. After this, the participants were asked about their general interest in art. Before the presentation of the paintings, they were able to test and adjust their speakers with a short audio-text. Each painting was introduced by a written instruction of the title and the name of the artist. The participants clicked the continue-button when they had read the information. After this they had to click the play-button to start the presentation of the painting together with the respective audio-text. Directly after the presentation of each painting, the participants were asked to report their experienced transportation, their aesthetic evaluation including liking, surprise, confusion and interest, and their subjective understanding of the painting. The presentation of the three paintings was done in random order to prevent order effects. After the presentation of all three paintings, the participants filled out their demographics, a question about prior knowledge of the paintings, and whether

they work or study in the field of art, history, or art-history. They were then debriefed and paid 4.50 £. The study received institutional research ethics committee approval.

Results

Control Variable: General Interest in Art

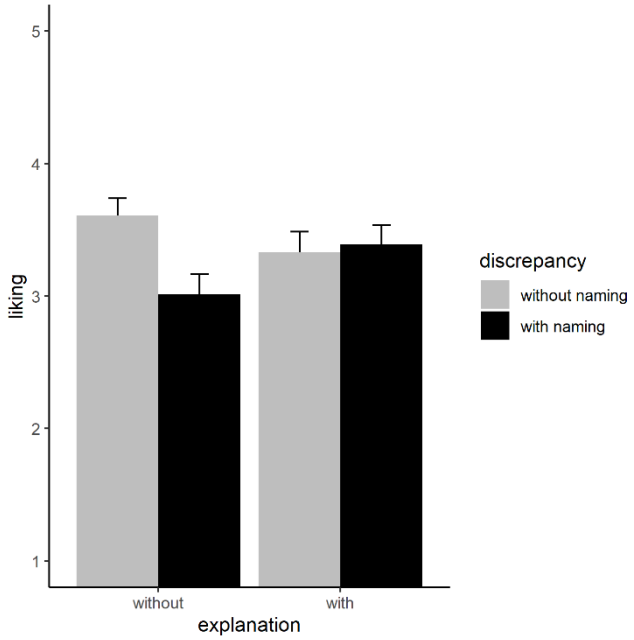
A one-way analysis of variance (ANOVA) was calculated with condition (I-E- vs. I+E- vs. I-E+ vs. I+E+) as the between-subjects factor. The analysis revealed no differences in general art interest between the four conditions, $F(3, 135) = 0.45$, $p = .721$, $\eta^2_p = .010$. Therefore, differences between the conditions cannot be explained by differences in general art interest.

Subjective Understanding

A two-way analysis of variance (ANOVA) was calculated across the three paintings with naming of inconsistencies (with vs. without) and an explanation of the intentions of the artist (with vs. without) as between-subjects factors. The analysis revealed no significant main effect of naming inconsistencies, $F(1, 135) = 0.79$, $p = .376$, $\eta^2_p = .006$. Subjective understanding did not differ significantly with inconsistencies named ($M = 5.27$, $SD = 0.99$) compared to without inconsistencies named ($M = 5.42$, $SD = 0.97$). The main effect of explanation was not significant, $F(1, 135) = 0.38$, $p = .540$, $\eta^2_p = .003$. Subjective understanding did not differ significantly with explanation ($M = 5.40$, $SD = 0.90$) compared to without explanation ($M = 5.32$, $SD = 1.05$). In contrast to our expectations, the two-way interaction between naming inconsistencies and explanation was also not significant, $F(1, 135) = 0.97$, $p = .326$, $\eta^2_p = .007$. The Bonferroni-adjusted comparison showed neither a significant difference between with inconsistencies named (I+E-: $M = 5.14$, $SD = 1.02$) and without inconsistencies named (I-E-: $M = 5.45$, $SD = 1.06$) when no explanation was given, $p = .186$, nor between with inconsistencies named (I+E+: $M = 5.41$, $SD = 0.96$) and without inconsistencies named (I-E+: $M = 5.39$, $SD = 0.88$) when an explanation was given, $p = .946$.

Figure 3

Interaction Between Discrepancy and Explanation With Regard to Liking.



Note. The error bars indicate standard errors.

Liking

A two-way analysis of variance (ANOVA) was calculated across the three paintings with naming of inconsistencies (with vs. without) and an explanation of the intentions of the artist (with vs. without) as between-subjects factors (see Figure 3). The analysis revealed no significant main effect of naming inconsistencies, $F(1, 135) = 3.23$, $p = .074$, $\eta^2_p = .023$. Liking did not differ significantly with inconsistencies named ($M = 3.20$, $SD = 0.82$) compared to without inconsistencies named ($M = 3.47$, $SD = 0.91$). The main effect of explanation was not significant, $F(1, 135) = 0.11$, $p = .736$, $\eta^2_p = .001$. Liking did not differ significantly with explanation ($M = 3.36$, $SD = 0.89$) compared to without explanation ($M = 3.35$, $SD = 0.88$). However, the two-way interaction between naming inconsistencies and explanation was significant, $F(1, 135) = 4.89$, $p = .029$, $\eta^2_p = .035$. As it was expected, the Bonferroni-adjusted comparison showed a significant lower liking with inconsistencies named (I+E-: $M = 3.01$, $SD = 0.84$) than without inconsistencies named (I-E-: $M = 3.61$, $SD = 0.83$) when no explanation was given, $p = .005$, but when an explanation was given, liking was equally high with (I+E-: *M*

= 3.39, *SD* = 0.77) and without inconsistencies named (I-E+: *M* = 3.33, *SD* = 0.98), *p* = .771. To check whether the effect was similar across the paintings, we calculated an additional ANOVA including painting as a within-factor. The ANOVA with the three factors inconsistencies named, explanation, and painting revealed no significant three-way interaction $F(2, 270) = 0.01$, $p = .994$, $\eta^2_p < .001$. The two-way interaction was therefore valid for all three paintings. Further we checked the linkage between liking, subjective understanding and other aesthetic emotions. Liking ratings correlated positively with subjective understanding ($r = .34$ $p < .001$), interest ($r = .71$, $p < .001$), surprise ($r = .54$ $p < .001$), and also transportation ($r = .65$ $p < .001$).

Surprise

A two-way analysis of variance (ANOVA) was calculated across the three paintings with naming of inconsistencies (with vs. without) and an explanation of the intentions of the artist (with vs. without) as the between-subjects factors. The analysis revealed no significant main effect of naming inconsistencies, $F(1, 135) = 0.18$, $p = .669$, $\eta^2_p = .001$. In contrast to our expectations, surprise did not differ with inconsistencies named (*M* = 2.37, *SD* = 0.82) compared to without inconsistencies named (*M* = 2.43, *SD* = 0.87). The main effect of explanation was not significant, $F(1, 135) = 0.43$, $p = .512$, $\eta^2_p = .003$. Surprise did not differ significantly with explanation (*M* = 2.35, *SD* = 0.80) compared to without explanation (*M* = 2.47, *SD* = 0.89). The two-way interaction between naming inconsistencies and explanation was not significant, $F(1, 135) = 1.37$, $p = .245$, $\eta^2_p = .010$. The Bonferroni-adjusted comparison showed neither a significant difference between with inconsistencies named (I+E-: *M* = 2.33, *SD* = 0.89) and without inconsistencies named (I-E-: *M* = 2.57, *SD* = 0.89) when no explanation was given, *p* = .258, nor between with inconsistencies named (I+E+: *M* = 2.41, *SD* = 0.76) and without inconsistencies named (I-E+: *M* = 2.30, *SD* = 0.84) when an explanation was given, *p* = .603.

Interest

A two-way analysis of variance (ANOVA) was calculated across the three paintings with naming of inconsistencies (with vs. without) and an explanation of the intentions of the artist (with vs. without) as the between-subjects factors. The analysis revealed no significant main effect of naming inconsistencies, $F(1, 135) = 0.46$, $p = .499$, $\eta^2_p = .003$. Interest did not differ significantly with inconsistencies named ($M = 3.57$, $SD = 0.90$) compared to without inconsistencies named ($M = 3.46$, $SD = 0.99$). The main effect of explanation was not significant, $F(1, 135) = 0.01$, $p = .913$, $\eta^2_p < .001$. Interest did not differ significantly with explanation ($M = 3.52$, $SD = 0.94$) compared to without explanation ($M = 3.52$, $SD = 0.94$). In contrast to our expectations, the two-way interaction between naming inconsistencies and explanation was not significant, $F(1, 135) = 0.57$, $p = .453$, $\eta^2_p = .004$. The Bonferroni-adjusted comparison showed neither a significant difference between with inconsistencies named (I+E-: $M = 3.39$, $SD = 1.06$) and without inconsistencies named (I-E-: $M = 3.62$, $SD = 0.85$) when no explanation was given, $p = .311$, nor between with inconsistencies named (I+E+: $M = 3.53$, $SD = 0.94$) and without inconsistencies named (I-E+: $M = 3.52$, $SD = 0.96$) when an explanation was given, $p = .958$.

Confusion

A two-way analysis of variance (ANOVA) was calculated across the three paintings with naming of inconsistencies (with vs. without) and an explanation of the intentions of the artist (with vs. without) as the between-subjects factors. The analysis revealed no significant main effect of naming inconsistencies, $F(1, 135) = 0.19$, $p = .661$, $\eta^2_p = .001$. Confusion did not differ significantly with inconsistencies named ($M = 1.64$, $SD = 0.59$) compared to without inconsistencies named ($M = 1.59$, $SD = 0.58$). The main effect of explanation was not significant, $F(1, 135) = 0.09$, $p = .760$, $\eta^2_p = .001$. Confusion did not differ significantly with explanation ($M = 1.60$, $SD = 0.54$) compared to without explanation ($M = 1.63$, $SD = 0.62$). In contrast to our expectations, the two-way interaction between naming inconsistencies and explanation was not significant, $F(1, 135) = 0.07$, $p = .792$, $\eta^2_p = .001$. The Bonferroni-adjusted comparison showed neither a significant difference between with inconsistencies

named (I+E-: *M* = 1.67, *SD* = 0.68) and without inconsistencies named (I-E-: *M* = 1.60, *SD* = 0.58) when no explanation was given, *p* = .975, nor between with inconsistencies named (I+E+: *M* = 1.61, *SD* = 0.50) and without inconsistencies named (I-E+: *M* = 1.59, *SD* = 0.58) when an explanation was given, *p* = .707.

Transportation

A two-way analysis of variance (ANOVA) was calculated across the three paintings with naming of inconsistencies (with vs. without) and an explanation of the intentions of the artist (with vs. without) as the between-subjects factors. The analysis revealed no significant main effect of naming inconsistencies, $F(1, 135) = 0.38$, $p = .540$, $\eta^2_p = .003$. Transportation did not differ significantly with inconsistencies named (*M* = 4.19, *SD* = 1.13) compared to without inconsistencies named (*M* = 4.31, *SD* = 1.17). The main effect of explanation was not significant, $F(1, 135) = 0.01$, $p = .917$, $\eta^2_p < .001$. Transportation did not differ significantly with explanation (*M* = 4.26, *SD* = 1.16) compared to without explanation (*M* = 4.25, *SD* = 1.15). In contrast to our expectations, the two-way interaction between naming inconsistencies and explanation was not significant, $F(1, 135) = 0.11$, $p = .739$, $\eta^2_p = .001$. The Bonferroni-adjusted comparison showed neither a significant difference between with inconsistencies named (I+E-: *M* = 4.14, *SD* = 1.15) and without inconsistencies named (I-E-: *M* = 4.33, *SD* = 1.15) when no explanation was given, *p* = .501, nor between with inconsistencies named I+E+: (*M* = 4.23, *SD* = 1.13) and without inconsistencies named (I-E+: *M* = 4.29, *SD* = 1.19) when an explanation was given, *p* = .844.

Discussion

Presently, evidence of the effects of additional information that is intended to foster meaning making on subjective understanding and liking of artworks is mixed (Chmiel & Schubert, 2019). Positive effects on liking are shown for highly abstract paintings but seldom for representational art (Leder et al., 2006; Moore & West, 2012; Swami, 2013). In two previous studies, negative effects on liking were found for unexplained mismatching titles, indicating that the processing of inconsistencies can lower the liking of abstract but also

representational artworks (Belke et al., 2010; Gerger & Leder, 2015). Following these findings, we presented historical paintings together with additional information either naming their historical inconsistencies or not. In addition, we either provided an explanation for the inconsistencies or not. Based on the present results and theories of art processing (VIMAP; Pelowski et al., 2017), we expected that the naming of a historical paintings' inconsistencies and leaving them unexplained would produce a similar effect like the unexplained mismatching titles, namely, a lower subjective understanding and lower liking, due to the viewer's processing of inconsistent information. In addition, we expected that the provision of an explanation would compensate for the negative effects of the naming of unexplained inconsistencies on subjective understanding (P1) and liking (P2).

Although subjective understanding was not significantly affected by the additional information in our experiment, liking was significantly lower when the inconsistencies of a painting were named but remained unexplained compared to when the inconsistencies were not just named, but also explained. In contrast, no similar difference in liking with or without explanations for accurate elements were found. Because liking under conditions of naming and explaining inconsistencies was similar to conditions of not naming inconsistencies, we conclude that the explanations did compensate for the detrimental effects of naming inconsistencies. Our result of a lower liking when inconsistencies are named but not explained is in line with the assumptions of the fluency theory, VIMAP, and previous empirical evidence (Belke et al., 2010; Gerger & Leder, 2015) and indicates that the negative effects in cases of missing explanations not only apply to semantically mismatching titles but also to longer explanations of representational art that require the viewer to process inconsistencies. Furthermore, our results show that the negative effects of unexplained inconsistencies on liking can be compensated for by an explanation of these inconsistencies. Hence, our study provides indication against the assumption that only the liking of abstract art can profit from additional information (Leder et al., 2006; Swami, 2013) and supports the claim that the reason for rare evidence of additional information affecting liking of

representational art might be a frequent ceiling effect resulting from the already high liking of representational art (Chmiel & Schubert, 2019).

Our results of a significant two-way interaction between inconsistency named and explanation for liking without a significant two-way interaction for subjective understanding seem to be more compatible with the VIMAP than with the psycho-historical framework for the science of art appreciation. While the latter emphasizes the role of understanding on the liking of artworks the former assumes that liking is a product of the congruency-check, which includes understanding, but also other factors, such as whether the painting matches the viewers expectations. Also, in accordance with the fluency theory (Reber et al., 2004), a less fluent processing remains to be a possible explanation, that needs to be examined more directly with additional processing measures in future studies. Fluency was also considered to be the underlying mechanism in the studies using unexplained mismatching titles (Belke et al., 2010; Gerger & Leder, 2015).

Regarding aesthetic emotions, we expected that surprise will be higher when inconsistencies are named than when they are not named, independent of whether these inconsistencies are explained or not (P3). We expected that interest will be lower when inconsistencies are named but unexplained than when they are not named and explained and that this effect will be compensated by an explanation for the inconsistencies (P4). On the contrary, we expected that confusion will be higher when inconsistencies are named and unexplained than when they are not named and explained, and that this effect is compensated again by an explanation for the inconsistencies (P5). In contrast to our expectations, we could not show any effects of naming the painting's inconsistencies on any of these emotional outcomes. Neither surprise nor interest were lowered, nor did confusion increase by informing the viewers about inconsistencies without explaining them. Subsequently, we could also not show a compensating effect of explanations for interest and confusion. Research on aesthetics often reported non-significant effects of additional information, such as titles, on emotional outcomes (Bubić et al., 2017) or more specifically

interest (Gerger & Leder, 2015; Leder et al., 2006). These results, however, are surprising since a close link between liking and the experience of aesthetic emotions can be assumed to be based on the models (Pelowski et al., 2017). Indeed, in our study, interest, surprise, and also transportation correlated significantly and highly positively with liking.

However, if our manipulation affected liking but neither subjective understanding nor aesthetic emotions, by which means was liking affected? In line with Gerger and Leder (2015) who found similar effects on liking but not on interest, we speculate that viewers based their lower liking on a greater disfluency in the condition in which inconsistencies were named and no explanation was given than in the conditions in which inconsistencies were not named. In addition, if inconsistencies were named but an explanation was given, the explanation may have reinstated fluency for subsequent processing and thereby compensated for the negative effect of naming unexplained inconsistencies on ratings of liking. This result is in line both with the fluency theory and the VIMAP. According to the VIMAP, this pattern would be expected if viewers base their liking judgements mainly on bottom-up processing (stage 2 to stage 4) such as fluency but do not engage in higher order cognitive processes. Because particularly laypersons tend to rely on lower stages of art processing for their evaluation of artworks (Mullennix & Robinet, 2018), we speculate that our manipulation affected art processing on lower stages but not on higher stages where understanding and aesthetic emotions would have been affected. Since audio explanations are intended to be used primarily by non-experts it would be interesting for future research to investigate whether and how audio explanations can also substantially affect laypersons higher order processing and thereby aesthetic emotions and the subjective understanding.

Based on theories of narrative processing (Busselle & Bilandzic, 2008), we expected lower transportation when unexplained inconsistencies are named than when they are not named, but this effect should be compensated for by an explanation (P6). We could not show that the naming of unexplained inconsistencies lowers transportation. Hence, our experiment provides no support for the model of narrative engagement (Busselle & Bilandzic, 2008) for

static pictorial narratives. Moreover, the effects on liking in our study are not comparable to effects on transportation being associated with enjoyment.

Some limitations must be noted: For aesthetics, we relied solely on self-reports. Physiological or other process measures, however, could be helpful as additional indicators of emotions and liking and the possibly less fluent processing. For spoken text, it might be important to consider not only what additional information is presented, but also how it is presented. In our study, we stressed the words similarly whether inconsistencies were presented or not. In a realistic context, surprising facts can be presented with a voice emphasizing this surprise which could foster effects of additional information on aesthetic emotions. Further, regarding emotions, it might be that the between-design of our study with participants either always viewing inaccurate paintings or participants always viewing accurate paintings prevented effects of the naming of unexplained inconsistencies. A within-design might be more suitable for investigating these effects on emotions. For example, Russell (2003) did not detect effects on aesthetic evaluation in the first experiment by using a between design but did in the second experiment by using a within design. Regarding the effects of an explanation of inconsistencies on liking, we always explained the inconsistencies by mentioning the good intentions of the artists. It would be interesting whether malevolent intentions would have similar effects or not. This could help to disentangle whether the effects result from the explanation itself or the additional positive information about the artist.

The reported effect sizes in previous studies for liking varied greatly in magnitude, depending on the information provided. Studies using titles as additional information often reported small effect sizes (Belke et al., 2010; Gerger & Leder, 2015), whereas high effect sizes are reported for content-specific information (Swami, 2013). As the additional information that we manipulated was a content specific interpretation and due to scarce previous studies on the effects of additional information on emotions, such as surprise and confusion, we decided to assume a medium effect of $f = 0.25$ for our study. Based on a

power analysis using G-Power, 128 participants were required for a power (1-β) of .80 to detect a medium effect of $f = 0.25$ and $\alpha = .05$. Due to our strict exclusion criteria, we had to exclude more participants than expected, and the remaining 139 participants were not equally distributed across the conditions, resulting in a sufficient but slightly lower power than we had aimed for. Regarding generalizability, we only considered laypersons of art. We would not expect the same results on liking for viewers more proficient in art and art-history due to their higher order processing or because they might be able to explain the inconsistencies themselves without the need for an external explanation by the audio-text (Bullot & Reber, 2013).

In conclusion, we could show that the naming of unexplained inconsistencies impairs the liking of representational paintings. However, an explanation about the inconsistency was able to compensate for this negative effect of additional information on the liking of representational artworks. Our results corroborate theories of art processing, such as the VIMAP (Pelowski et al., 2017), and show that not only abstract art can profit from additional information but also representational art. Our results extend the present literature by showing that negative effects of additional information hold not only for unexplained mismatching titles (Belke et al., 2015; Gerger & Leder, 2015) but also for informing about the inconsistencies of the content of a painting with regard to reality and at the same time leaving these inconsistencies unexplained. In contrast to unexplained mismatching titles, information about a painting's historical inconsistencies is frequently provided in museums of art as this is an important part of the interpretation of a painting's content. Therefore, our results have practical implications for the design of information accompanying representational artworks in museums. First, additional information can not only enhance the liking of artworks but also lower the liking of artworks if it requires a layperson to process unexplained inconsistencies. Second, if inconsistencies of a representational painting are first named unexplained, the liking can be restored when an explanation for the inconsistencies is added in a second step.

Declaration according to § 5 Abs. 2 No. 8 of the PhD regulations of the Faculty of Science

-Collaborative Publications-

The following chapter (Chapter 4) consists of a manuscript that was authored by me and is not published.

Author	Author position	Scientific ideas (%)	Data generation (%)	Analysis and Interpretation (%)	Paper writing (%)
Manuel Knoos	First Author	100	100	100	100
Title of the paper	Conflict Regulation Strategies with Multiple Documents of Conflicting Pictures of Historic Events				
Status in publication process:	Not published				

4. Study 4

Conflict Regulation Strategies with Multiple Documents of Conflicting Pictures of Historic Events

For a deep understanding about a complex topic, such as historic events, climate change, or political events, humans need to incorporate content from multiple documents providing different perspectives on the topic. Thereby the content of a document and the information about the authors must be considered together (Wineburg, 1991), because often the authors make one sided claims and provide conflicting perspectives. This is similarly true for texts and pictures (Burke, 2001). In many situations, humans rely on pictorial documents to acquire information about a topic. For example, photography's in newspapers, TV news, commercials, social media, archeological reconstructions or museum objects. In many of these situations the acquisition of information from the pictures is supported by written or spoken text but the picture is the main document providing evidence for claims about the topic. For climate change this can be graphs plotting the worlds carbon dioxide output together with mean temperature over the years with a sharp increase since the industrialization. While authors on the other side might show plots including longer periods and point out the volatility throughout the ages. For a historical topic these might be two different paintings, or photography depicting the same event differently. This situation of information acquisition with two or more documents can be understood within the framework of multiple documents (e.g. Britt & Rouet, 2012). Presently, multiple document situations are mostly investigated with text documents and research on pictorial documents is neglected. In our study we therefore investigated multiple picture documents providing conflicting information about a historic event together with an audio-text supporting information acquisition from the pictures.

Learning with Multiple Conflicting Text Documents

The framework of multiple documents (Britt & Rouet, 2012) assumes that readers of multiple documents first construct a mental representations about the content of each

document. These separate mental representations are then integrated into one coherent mental model about the topic. Readers also store information about the authors in document nodes and link these nodes with each other and to the content of the respective documents in so called intertext predicates. These intertext predicates allow the reader to construct a coherent mental model by attributing conflicts to different author's opinions. Attending author information, evaluating, representing, and using it is referred to as sourcing. Research shows that learners engage in sourcing more often when the documents provide conflicting information than when the documents provide consistent information (Braasch et al., 2012; Braasch & Bråten, 2017; Bråten & Braasch, 2018).

Sourcing is only one of the strategies applied for the regulation of conflicts. The content-source-integration model (CSI; Stadtler & Bromme, 2014) describes the readers' process of conflict regulation in more detail by proposing three stages of conflict regulation. First, readers need to detect the conflict. Second, readers need to regulate the conflict and restore coherence. To restore coherence, readers can either ignore the conflict, which means they detect the conflict but do not remember it afterwards, or they reconcile the conflicting claims, or they draw inferences, or lastly they can accept the conflict due to different sources. In the third stage, readers can try to resolve the conflict between two authors by either asking what is true or if this is not possibly answered, they can resolve the conflict by asking whom to believe. When asking whom to believe the readers need to judge the trustworthiness of a document and its author (cf. Saux et al., 2018). Thereby they consider expertise or the benevolence of the involved authors (Thomm & Bromme, 2016). In the recipient's perspective, a benevolent author intends to provide the best possible information to the recipient instead of following selfish interests or being biased by other goals, such as commercial interests (Mascaro & Sperber, 2009; Thomm & Bromme, 2016). In line with the theory, research shows that conflict lead to longer processing of the author information when the authors differ in benevolence than if they do not differ in benevolence (Gottschling et al., 2019). Participants better remembered the authors names when expertise or trustworthiness differences between the authors were present compared to when they

were absent (Thomm & Bromme, 2016). In addition, information about the authors that is relevant for the conflict is remembered better than irrelevant information, such as visual appearance of the authors (Saux et al., 2018). Arguments from benevolent authors are rated more useful than commercially biased authors (Kammerer et al., 2016) and trustworthy authors are cited more often when readers are asked to write an essay about the topic (List, Alexander, & Stephens, 2017).

Besides the resolution mechanisms described by the CSI-model, readers can postpone the conflict resolution if the given information is not sufficient in order to resolve the conflict to an acceptable degree for the reader (Graesser et al., 1994) and become more interested in receiving further information to resolve the conflict. This assumption is corroborated by readers asking more information seeking questions after encountering conflicting information (Graesser & Olde, 2003; Otero & Graesser, 2001).

Learning with Multiple Documents of Different Media Formats

Analogous to the surface-, text-base-, and the situational model of the Construction-Integration model (Kintsch, 1988) describing text processing, viewers of pictures construct representations on three different levels (Solso, 2003). These three levels proposed by Solso (2003) are similar to the representations proposed by Kintsch (1988) (Millis & Larson, 2008). Furthermore, learners use mostly similar learning strategies with texts and pictures (Loughlin et al., 2015) and also with multiple documents of text or videos (List, 2018). One study (Salmerón et al., 2020) used conflicting documents of the same or different media formats. Thereby primary school students either learned with two text documents or with two video documents or one video and one text document. The results indicated no differences in source use measured as citations in a subsequent essay. However, students better remembered the profession of one author when it was visibly presented in the video than when it was only presented textual. In contrast, integration of information was better with texts (Salmerón et al., 2020). Better integration with texts than videos was also found by List and Ballenger (2019). One reason for better integration might be that the learners' multiple

document competences are mostly trained with text documents in (Britt & Aglinskas, 2002) whereas pictures are tendentially neglected at least in history education (Burke, 2001). Especially, sourcing is a competence applied frequently by experts but less often by laypeople (Wineburg, 1991). Therefore, it is unclear whether recipients are able to use the information about sources of multiple pictures when it is useful for resolving conflicts in the same way as they would do for texts.

Present Study

In the present study we used paintings of artists making conflicting claims about the same historic event. Our goal was to investigate how the presence or absence of differences in the artist's benevolence affects source processing, evaluation, use and memory for information that is relevant for the conflicts. We manipulated the information about the artists' benevolence in our study by following the design of multiple text studies (e.g. Gottschling et al., 2019). Different from previous text studies, we did not rely on professions associated with higher or lower benevolence (e.g. researcher in industry vs. professor at a university) but manipulated benevolence more directly. For the condition with differences in benevolence between the artists, we described one artist with a documentary intention and one artist with a propagandistic intention. The comparison condition consisted of two sub-groups. For one sub-group both artists were described with a propagandistic intention (without difference) and the other sub-group received no information about the artists intentions at all (without intentions named). Based on the theory, we expected differences between the conditions with differences in benevolence and the two subgroups taken together. We expected no differences between the two subgroups (without differences in benevolence vs. without intentions named), as the source information is not useful to resolve the conflict in these groups.

We hypothesized that (1) with differences in benevolence compared to without differences and without intentions named, the source processing is enhanced during a task following the audiovisual presentation, which means that viewers click more often on the

button to read the information again and process this information longer. (2) With differences in benevolence compared to without differences and without intentions named, the processing times for the two paintings differ more greatly during a task following the audiovisual presentation. (3) With differences in benevolence compared to without differences and without intentions named, the message credibility of the two paintings differ more greatly. (4) With differences in benevolence compared to without differences and without intentions named, the sources are more often used and cited in a subsequent essay. (5) With differences in benevolence compared to without differences and without intentions named, the memory about the artists name and information relevant for the conflict is better. (6) With differences in benevolence compared to without differences and without intentions named, the interest in further information about the conflicting pictorial element and the artists expertise is lower.

Method

Participants

228 Participants were recruited on Prolific. The online experiment was done with the software IWM-Study and could be accessed with all common browsers. Participants were requested to participate only via computers and neither with mobile phones nor tablets, due to the small screen size. Participants were pre-filtered to include only German citizens and native speakers of German. From the 228 participants, two were excluded because they refused the use of their data for our analysis. 15 were excluded because they reported that they used search engines during the experiment or experienced longer interruptions during the experiment. 36 were excluded because they already knew at least one of the four paintings. 8 were excluded because they had higher than medium pre-knowledge about the artists or the historic events. Three were excluded because they participated via smartphone. Two were excluded due to technical issues resulting in missing data. One reported note taking during the presentation of the paintings and was excluded. Finally, 32 participants were excluded because their viewing times of the artist information was shorter than 10

seconds indicating that they did not read the whole information about the artists. Subsequently, 129 participants remained for the analysis: 48 (38%) females, 80 (63%) males, one diverse (< 1%); aged between 18 and 60 years (M = 28.44, SD = 8.19).

Design

We tested our hypotheses using a 3x1 design with intentions of the artists (with difference in intentions vs. without difference in intentions vs. without intentions named) as the between-subjects factors. The 129 participants were randomly assigned to one of our three conditions (with difference in intentions n = 42, without difference in intentions n = 41, without naming any intentions n = 46). The study was approved by the institution's ethical committee.

Materials

The material consisted of two pairs of historical paintings. The paintings of the first pair were created by Emanuel Leutze and George Caleb Bingham and are both entitled "Washington crossing the Delaware". They depict George Washington on a boat crossing the Delaware River together with his soldiers during the American Revolution. The paintings of the second pair were created by Benjamin West and Edward Penny and are both entitled "The Death of General Wolfe". They depict the dying British General Wolfe surrounded by his soldiers during the Battle of Quebec in Canada. We chose these paintings because the two paintings of a pair contain at least two similar pictorial elements indicating the same fact about the historical event, and two pictorial elements that are depicted differently indicating conflicting facts about the historic event. For example, a musket was depicted in both paintings of General Wolfe, whereas West's painting depicts a Scottish soldier pointing to the soldier carrying the message of their victory and Penny's painting depicts a British grenadier pointing to the soldier carrying the message of their victory.

For each picture pair, we created a video explaining the two paintings. In the beginning of the videos, general information was given about the paintings, the event and the

main character. During this part both paintings were visible in the video. After this the video presented only one painting and explained one pictorial element which was visually signaled by greying out all other pictorial elements to support identification. Then the video switched to the other painting of the pair and explained the corresponding pictorial element of this painting. In this way, first the two pictorial elements where described indicating the same fact with regard to the historic event. Then two elements were described indicating conflicting facts. At the end of the video again both pictures were shown, and the outcome of the event was described briefly. The content of the videos was the same in all three conditions, but we counterbalanced the order of the presentation of the paintings to control for order effects. For about half of the participants in each condition the video started first describing a pictorial element of picture A and then picture B and so on for all four pictorial elements. For about the other half of the participants, this was done oppositely starting with picture B. The video for the picture pair of Leutze and Bingham was 3:01 minutes long and the video for the picture pair of West and Penny was 3:09 minutes long.

In addition, we created written texts informing about each of the four artists (see Appendix for an example). The first part of this text described some general biographical stations in the artists life. The second part was about the painting presented in our study which we described as one of the artist's most important works. The second part included information about the year of production and an event that took place during the time the artist created the painting. Then the last sentence named the audience which the artist had in mind and manipulated his intention for the painting according to the three conditions. Either one of the two artists of each pair were described with a propagandistic- and the other with a documentary intention, or both were described with a propagandistic intention, or a text was provided not revealing any intentions for both artists. When one artist was named as propagandistic and the other as documentary, we controlled for effects of the different artists and their respective paintings by counterbalancing which artist's intention was named as propagandistic and which as being documentary.

Measures

As a manipulation check we measured source credibility with the 6-Item questionnaire (Flanagin & Metzger, 2003). Participants rated the artists on trustworthiness, believability, reliability, authoritativeness, honesty and bias each on a 7-point Likert-Scale ranging from 1 (not at all) to 7 (very much). Means were calculated for each artist. We subtracted the source credibility of the two artists of each pair from each other. Then we calculated the mean of the absolute values of differences between the two artists of each pair.

We measured sourcing as clicks on the icons of the artist information and as viewing times of the information about the artists, after the video, when participants were able to freely view the paintings and the information about the artists again. We calculated sum scores for clicks on artist. We calculated the sum of viewing times for each time the participant viewed the information after the video. Then we calculated the mean score for viewing times of all four artists.

We measured picture processing as the differences in viewing times of the two paintings of each pair. We summed up the viewing times for each painting each time the painting was viewed after the video. Then we subtracted the sum of viewing times of the two paintings of each pair from each other. Lastly, we calculated the mean of the absolute values of differences between the two paintings of each pair.

We measured message credibility with the 5-Item questionnaire (Flanagin & Metzger, 2003). Participants rated the paintings on trustworthiness, completeness, believability, bias and accuracy each on a 7-point Likert-Scale ranging from 1 (not at all) to 7 (very much). For each participant we calculated the mean for each painting. We subtracted the participant's message credibility mean of the two paintings from each other for each pair. Then we calculated the mean of the absolute values of differences between the two paintings of the two pair.

We measured source use as the artists name cited in the participants' essays. Participants received one point for each name of the four artists given in their essay then we calculated the mean. As we expected spelling or typing errors in the viewers essays, names highly similar to the artists real name, such as "Leutz" or "Leuze" were accepted as a correct answer and received one point.

For the memory about the source and information relevant for the conflict, we asked about the artists name, the audience intended by the artist and the event that took place during the production of the painting. Participants received one point for each correctly remembered fact. Again, we expected spelling or typing errors in the viewers responses and answers highly similar to the artists real name or the name of the event or audience were accepted as a correct answer and received one point. We calculated the mean for all facts about the artists.

We measured interest about the conflicting pictorial elements by asking how much the participants are interested in receiving further information about the conflicting pictorial elements. The interest for the two conflicting pictorial elements of each pair was rated on a 7-point Likert-Scale ranging from 1 (not at all) to 7 (very much). We calculated the mean score. Interest in further information about the artists expertise was also measured with two items. Participants rated their interest about the artists expertise and about the sources the artists considered for the creation of their paintings on a 7-point Likert-Scale ranging from 1 (not at all) to 7 (very much). We subsequently calculated the mean score.

Procedure

At the beginning of the experiment, participants were informed about the study and filled in the consent. They were able to test and adjust their speakers using a short audio-text. Then questions about pre-knowledge followed asking how much the participants already knew about the four artists and the two events presented in our study. After this, the participants were instructed that they will read information about artists and later see the paintings of these artists in a video. After the video they will be able to freely view the

material again in order to solve a task that is presented directly after the video. In addition to this task, they will have to write a short essay about the historic event using the information given in the audiotext and the paintings of the video. After the instruction, the first pair of artists were announced by stating that the subsequently presented artists created a painting about Washington's crossing of the Delaware River. Then participants were able to read the information about the two artists sequentially. The two texts about the artists were presented in a counterbalanced order following the order of the paintings in the video. When the video started with presenting the first pictorial element form painting A then information about artist A was also given first. After reading the artist information participants were able to view the video. When the video was finished participants were instructed with a task. For the paintings about Washington the task was to inspect the material again in order to describe how Washington eventually looked like during the event. Participants were able to navigate through the material consisting of the information about the artists and their paintings by clicking on small icons of the documents located in the bottom of the screen. When they were finished, they clicked on the button on the upper right to continue with writing down a solution for the tasks as well as writing a short essay about the historic event. After this the presentation of the second pair of artists and their respecting paintings followed in the same structure of the first pair. The task for the second pair about the death of General Wolfe was to describe the environment of the battlefield.

When participants had finished with videos and tasks for both pairs, they were asked to rate their trust in the artists and in the message credibility about the paintings. This was again done in the same counterbalanced order as the presentation of the paintings and the text about the artists. After this they rated their interest in receiving further information about the consistent pictorial elements, the conflicting pictorial elements and the expertise of both artists. Then the memory test for information about the artist and information relevant for the conflict followed again in the same counterbalanced order as the presentation of the paintings. Lastly participants answered demographic questions and were debriefed. They received 3.75 £. The study received institutional research ethics committee approval.

Results

We calculated one-way ANOVAs for all our dependent variables (see table 6 for descriptive statistics) together with two contrasts. The first contrast compared the condition with differences in intentions with the other two conditions taken together (without differences in intentions and without naming any intentions). The second contrast compared the conditions without differences in intentions and without naming any intentions. We considered the analysis as supporting our hypotheses if the overall ANOVA was significant as well as the first contrast. However, in line with our hypotheses, the second contrast should not be significant.

Table 6

Descriptive Statistics for all Dependent Variables

Dependent variable	Condition		
	With difference	Without difference	Without intentions
Difference in artist trust	0.80 (0.75)	0.61 (0.57)	0.49 (0.61)
Sourcing (clicks)	2.45 (2.41)	2.61 (2.56)	3.04 (2.65)
Source processing	3689.01 (5141.85)	3465.13 (4501.54)	6309.54 (7987.02)
Difference in picture processing	5541.43 (4090.05)	7631.90 (7148.92)	6295.36 (5764.96)
Difference in message credibility	0.85 (0.76)	0.89 (0.89)	0.82 (0.82)
Source use	0.11 (0.25)	0.10 (0.27)	0.12 (0.27)
Source memory	0.56 (0.24)	0.58 (0.24)	0.55 (0.24)
Interest in expertise	4.79 (1.62)	5.25 (1.33)	4.89 (1.63)
Interest in conflicting elements	3.44 (1.44)	3.59 (1.28)	3.34 (1.51)

The ANOVA for the manipulation check, namely the differences in trustworthiness of the artists, revealed no differences between the three conditions, $F(2, 126) = 2.45$, $p = .090$. However, the first contrast was significant, $t(126) = 2.03$, $p = .022$, revealing a significant difference between the condition with intentions ($M = 0.80$, $SD = 0.75$) and the two other conditions taken together ($M = 0.55$, $SD = 0.59$). The second contrast was not significant, $t(126) = 0.81$, $p = .416$, indicating no difference between the condition without differences in intentions and the condition without any intentions named.

The ANOVA for source processing regarding clicks on the buttons to read the source information revealed no differences between the three conditions, $F(2, 126) = 0.64$, $p = .528$. The first contrast was not significant, $t(126) = 0.78$, $p = .218$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to clicks on the button of the sources. The second contrast was not significant, $t(126) = 0.79$, $p = .429$, indicating no difference between the condition without differences in intentions and the condition without any intentions named with regard to clicks on the button of the sources.

For the analysis of source processing regarding viewing times of the artist information, 6 participants were additionally excluded from the original sample because their values were greater than three standard deviations away from the mean. The ANOVA for source processing regarding viewing times revealed no significant differences between the three conditions, $F(2, 120) = 2.80$, $p = .065$. The first contrast was not significant, $t(120) = 1.01$, $p = .314$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to source processing times. However, the second contrast was significant, $t(120) = 2.10$, $p = .038$, indicating that source processing times were shorter in the condition without differences in intentions ($M = 3465.13$, $SD = 4501.54$) than in the condition without any intentions named ($M = 6309.54$, $SD = 7987.02$).

For the analysis of picture processing times, 14 participants were additionally excluded from the original sample. 8 of them did not view any of the two pictures again after

the video, which indicates that they solved the task out of their memory. 6 of them because their values were greater than three standard deviations away from the mean. The ANOVA for picture processing times revealed no significant differences between the three conditions, $F(2, 112) = 1.27$, $p = .286$. The first contrast was not significant, $t(112) = 1.24$, $p = .110$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to picture processing times. The second contrast was not significant, $t(112) = 1.01$, $p = .315$, indicating no difference between the condition without differences in intentions and the condition without any intentions named with regard to picture processing times.

The ANOVA for message credibility revealed no significant differences between the three conditions, $F(2, 126) = 0.09$, $p = .913$. The first contrast was not significant, $t(126) = 0.05$, $p = .481$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to message credibility. The second contrast was not significant, $t(126) = 0.43$, $p = .671$, indicating no difference between the condition without differences in intentions and the condition without any intentions named with regard to message credibility.

The ANOVA for source use revealed no significant differences between the three conditions, $F(2, 126) = 0.06$, $p = .947$. The first contrast was not significant, $t(126) = 0.09$, $p = .466$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to citations in the subsequent essay. The second contrast was not significant, $t(126) = 0.32$, $p = .748$, indicating no difference between the condition without differences in intentions and the condition without any intentions named with regard to citations in the subsequent essay.

The ANOVA for source memory revealed no significant differences between the three conditions, $F(2, 126) = 0.12$, $p = .886$. The first contrast was not significant, $t(126) = 0.23$, $p = .411$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to source memory. The second contrast was not

significant, $t(126) = 0.44$, $p = .658$, indicating no difference between the condition without differences in intentions and the condition without any intentions named with regard to source memory.

The ANOVA for interest in further information about the artists expertise revealed no significant differences between the three conditions, $F(2, 126) = 1.03$, $p = .359$. The first contrast was not significant, $t(126) = 0.96$, $p = .171$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to interest in further information about the artists expertise. The second contrast was not significant, $t(126) = 1.10$, $p = .272$, indicating no difference between the condition without differences in intentions and the condition without any intentions named with regard to interest in further information about the artists expertise.

The ANOVA for interest in further information about the conflicting pictorial elements revealed no significant differences between the three conditions, $F(2, 126) = 0.35$, $p = .704$. The first contrast was not significant, $t(126) = 0.09$, $p = .475$, indicating no difference between the condition with intentions and the two other conditions taken together with regard to interest in further information about the conflicting pictorial elements. The second contrast was not significant, $t(126) = 0.84$, $p = .404$, indicating no difference between the condition without differences in intentions and the condition without any intentions named with regard to interest in further information about the conflicting pictorial elements.

Discussion

In our study we investigated multiple conflicting documents using picture material. We were interested in effects of the presence or absence of differences in the artist's benevolence on source processing and document processing, evaluation, use and memory for source information when learners are confronted with multiple conflicting picture documents. We expected that effects known from studies using text material (Gottschling et al., 2019; Kammerer et al., 2016; Saux et al., 2018) generalize to picture material.

In contrast to our expectations, the results of our study revealed no significant differences in sourcing between the condition with differences in artists' benevolence and the two condition without differences in artists' benevolence and without intentions named. Neither did participants click more often on the button for reading the source information again after the conflict nor did they process the information about the sources longer. Our results regarding sourcing are surprising as this is a strategy frequently applied by readers after encountering conflicts between two texts (Bråten & Braasch, 2018). A previous text study manipulating the benevolence of the sources of the conflicting documents also observed longer source processing when there were differences in benevolence than when there were no differences (Gottschling et al., 2019). In addition, in our study, the difference between processing times of the two picture documents was not higher with differences in benevolence compared to no difference and no intentions named. Hence, recipients did not use source information to focus their processing on the more trustworthy paintings to acquire reliable information for solving the task. According to the literature, documents are considered as being more useful when their sources appear trustworthy compared to documents of sources that are considered as less trustworthy (Kammerer et al., 2016). However, in our study we could not show that a potential higher usefulness rating of trustworthy documents, does result in higher use of the document for solving a task on the actual behavioral level. One explanation for this is a further result of our study, namely that the message credibility was not affected by our manipulation. There was no significant difference with regard to trustworthiness differences between the two pictures of each pair in the three conditions. This implies that although the participants in the condition with differences in benevolence of the artists perceived a greater differences between the trustworthiness of the two sources than the participants in the other two conditions, they did not perceive a greater difference in trustworthiness of the two picture documents created by these sources. In our study, participants did also not cite the artists of the paintings more often in a subsequent essay when sources differed in their benevolence than when they did not differ, or when no intentions were named. This result is surprising as the CSI model

(Stadtler & Bromme, 2014) and the document model (Britt & Rouet, 2012) proposes that recipients use source information and tag conflicting information to different sources in order to resolve conflicts. This should be enhanced if the source information can help answering the question of whom to believe (Stadtler & Bromme, 2014). In our study, we could not show better memory for information about the source in the condition with differences in artists' benevolence compared to the condition without differences in artists' benevolence and without intentions named. An enhanced memory for source information due to conflict is also a well-supported effect in text research (Braasch et al., 2012; Bråten et al., 2016; Saux et al., 2018). However, our result supports the findings of a previous text study manipulating benevolence of the sources. In this study source memory measured by the name of the source was not higher with differences in benevolence compared to no differences in benevolence of the two sources. As the overall memory for sources was high, the authors suggested that this could have been a potential ceiling effect (Gottschling et al., 2019). For this reason, we measured source memory finer grained by not only using the name of the sources but also asking for the memory about two facts relevant for the conflict. This was inspired by previous studies reporting that especially information relevant for the conflict is remembered better by readers (Saux et al., 2018). Although we used this finer grained measurement, we did not find significant differences in source memory. Therefore, based on our findings and present literature, conflicts might affect source memory independently whether the source information is useful to decide whom to believe (differences in benevolence of the sources) or not (no differences in benevolence of the sources). Our study further indicates that information seeking was no alternative conflict regulation strategy in the conditions where sourcing was not a useful strategy, namely in the condition without differences in benevolence and the condition where no intentions were named. Our results showed no difference with regard to interest, neither in further information about the artist's expertise nor about the conflicting pictorial elements, between the condition of differences in artists' benevolence and the condition without differences in artists' benevolence.

In sum, the results of our study do not support that differences in benevolence of sources affect the conflict regulation strategies that are applied by the recipient during learning with conflicting picture documents. This stands in fundamental contrast to studies using multiple text documents (Gottschling et al., 2019; Thomm & Bromme, 2016) and theories of multiple document comprehension (Britt & Rouet, 2012; Stadtler & Bromme, 2014). Our result can hint at potential differences between conflict regulation strategies when using multiple conflicting texts or pictures. When looking exploratorily at the memory for only the names of the paintings' sources in our study, our data reveals that viewers did only remember 56% of the painting's sources correctly. This is surprising as the name of each artist was stated in the written text before viewing the paintings and each artist was named 5 times during the video about the paintings. Information about the artists could also be accessed again after the video. In addition, our experiment only consisted of four paintings and their respective artists. The previous text studies encountered ceiling effects for memory about the name of the sources (Gottschling et al., 2019). That the viewers in our experiment did not construct reliable source-document links could explain the overall pattern of the results. One reason for this difference between previous text and our study using pictures might be a lack of training in formal education for using picture documents compared to text documents. Therefore, viewers might be inexperienced in evaluating source characteristics in order to resolve a conflict. Viewers might have considered information about the artists as less relevant and did not construct reliable source-document nodes. Another reason for our results differing from previous studies and the assumptions of theories is indicated by the manipulation check. Although the manipulation check regarding trustworthiness differences of the sources revealed a significant contrast between the condition with differences in benevolence and the two other conditions taken together, this difference was rather small, and the overall ANOVA was not significant. We used paintings commonly known as artworks and it might be that our manipulation of artists intending either a documentary purpose or a propagandistic purpose was not fully believed by the viewers. Therefore, for future studies we recommend using other pictures than history paintings, such as photographies. Thereby

either propaganda- or documentary photographies can be presented to the viewers. Also different pictures of consumer products with their sources either named as companies selling the product or instutions of consumer protection could be considered. For such pictures, the different intentions of sources might be more believable, and recipients might be more experienced of evaluating and using source information when considering pictures of product recommendations or other advertisement.

As always, some limitations of our study must be noted. First, due do the strict exclusion criteria our sample size was lower than we planned. This resulted in a lower power to detect the effects we expected. Second, we did not ask participants whether they remembered the intentions of the artists, but only checked our manipulation more inderectly by asking for source trust. Lastly, the manipulation in our study differed from the previous studies using text material. Text studies mostly used authors of different professions, such as a researcher in industry versus a researcher at a university (Gottschling et al., 2019; Kammerer et al., 2016; Thomm & Bromme, 2016). We manipulated the trustworthiness more directly by proposing different intention of the sources, namely propaganda and documentation. This might have been less authentic and believable and might not be encountered in this way in a realistic multiple document situation.

Despite the null effects and the limitations of our study, the results can be of interest to the field of multiple documents research. Our study indicates that a generalization of conflict regulation strategies from text to picture material is not trivial. Although pictures and texts are similarly represented (Millis & Larson, 2008) and processed with mostly similar strategies (Loughlin et al., 2015), the conflict regulation strategies, such as sourcing might differ. One reason could be different training experience with documents of text and picture. Sourcing is mostly trained with text documents and pictures are widely ignored at least in history education (Burke, 2001). Previous studies using video material showed that viewers were able to identify differences in the trustworthiness of sources presented in these videos but integration of information was worse than with texts (Salmerón et al., 2020). This is in line

with our arguments that there could be a lack of training as videos are far more common as learning material than static images but not as common as texts. However, since this study is to the best of our knowledge the first study using two picture documents, more research is needed to overcome limitations of a single study and to draw final conclusions about conflict regulation with picture documents.

5. General Discussion

Information is not only acquired and communicated in texts but frequently in the form of pictures. Analogous to texts viewers of pictures need to consider content together with information about the sources to avoid being seized by one sided perspective and distortions (Wineburg, 1991). Furthermore, audio-texts can help acquisition of knowledge from pictures also by directing attention and by enabling dual coding (Glaser et al., 2020; Glaser & Schwan, 2015). However, presenting pictures together with audio-texts naming the distortions of a picture creates a multiple document situation confronting the recipient with conflicting information about the historic event. In the present studies I investigated how conflicting information affects the viewers processing, memory and evaluation of pictures with regard to three aspects of the pictures. First, pictures as historical documents. In this case the pictures together with audio-texts are used as documents to acquire knowledge relevant for learning history (Burke, 2001). Second, pictures as narratives. Pictures about historic events often tell a story with their depiction including main characters and a plot. Recipients tendentially process stories in a narrative processing mode instead of an analytic mode. While the former emphasizes the role of emotions and a vivid experience of the facts as if oneself is in the world described by the narrative, the latter incorporates a more critical perspective on the presented facts (Green & Brock, 2000). Third, the aesthetic experience. This was a valuable perspective, since I used history paintings as picture material which are often exhibited in museums and inherit cultural value as artworks. Models of art processing make claims about how additional information can affect the art experience (Bullot & Reber, 2013; Pelowski et al., 2017; Reber et al., 2004). In the following, I will briefly present the results of the four studies before discussing the results with regard to theories and previous research in the context of the three aspects.

Study 1 revealed that recipients detected conflicts between audio-text and picture documents. This was indicated by longer fixation times on pictorial content when it was named with discrepancies than without discrepancies to the historic event. Regarding conflict

regulation strategies, recipients did neither process the information about the source longer nor report higher interest in further information with discrepancies named than without discrepancies named. Hence, results of Study 1 neither support sourcing nor information seeking as conflict regulation strategies. Lastly, the viewers' perceived transportation was lower when discrepancies were named compared to without the naming of discrepancies.

In Study 2, recipients rated the trustworthiness of the picture document lower when discrepancies were named but not explained, compared to not naming discrepancies. As expected, providing an explanation of the benevolent intentions of the source compensated the negative effect of naming discrepancies. This indicates that viewers engaged in sourcing and used the information about the source together with information about the content in order to rate the trustworthiness of the picture document. Memory for the discrepant pictorial content was not better with discrepancies named than without, but an exploratory analysis revealed that memory for the consistent pictorial content was worse when discrepancies were named. In Study 2, the effect of a lower transportation due to discrepancies was only replicated for one of three paintings and only when no explanation was provided.

In Study 3, recipients did not rate subjective understanding lower when discrepancies were named and not explained compared to not naming discrepancies. Subsequently an explanation could not compensate this effect. However, liking was significantly lower when discrepancies were named and not explained compared to not naming discrepancies. In line with our expectations, this effect was compensated by an explanation for the discrepancies. There was neither no significant effect of discrepancies and explanation on surprise, interest, confusion, and transportation.

In Study 4, recipients did not click on the button more often to read the source information again when differences in intentions of the artists were present (propaganda vs. documentary) compared to the group where differences were absent. Recipients did not process the information about the sources longer when differences in authors' benevolence were present than when they are absent. The differences in processing time of the two

paintings in a task following the presentation did not differ when differences in authors' benevolence were present than when they are absent. The differences in trustworthiness rating of the two paintings did not differ between when differences in authors' benevolence were present than when they are absent. Participants did not cite the source more often in a subsequent essay, their memory for source information was not better, and they did not report lower interest in further information about the source or conflicting elements when differences in authors' benevolence were present than when they are absent. Hence, in Study 4, participants did not engage in more sourcing and less information seeking when source information was useful for conflict regulation than when source information was not useful for conflict regulation.

5.1 Results Regarding Pictures as Historical Documents

In three of four studies I addressed the first aspect of pictures, namely the picture as a historical document (Study 1, Study 2, Study 4). In Study 1 and 2, viewers were confronted with conflicts between the depictions of historic events in the form of history paintings and accompanying audio-texts claiming how the depicted historic events "really" were according to today's historians' opinions. Thereby the audio-text directly commented on the picture document. This situation is very different from previous studies using mostly written text documents not directly commenting on each other. Therefore, I was interested whether effects known from multiple text studies generalize to such different material.

As expected, the results of Study 1 revealed higher fixation times on conflicting pictorial elements after the audio-text presented the pictorial elements as being discrepant with the real historic event than when they were presented without naming the discrepancies. These longer fixation times indicate that recipients noticed the conflict and possibly engaged in conflict regulation strategies. The result of longer processing due to conflicts is consistent with previous research. Conflicting information leading to longer processing times is a well-established effect in research on single text documents (Albrecht & O'Brien, 1993; O'Brien et al., 1998; O'Brien & Myers, 1985). Longer processing times on conflicting content compared

to non-conflicting content was also shown in a recent study using conflicts between text and picture on a single document page (Schüler, 2017) and on different pages of one document (Schüler, 2019). Since, participants detected conflicts they might have also tried to regulate and reconcile the conflict. In Study 1, I investigated source processing as a conflict regulation strategy indicated by fixation times on the source. In contrast to my expectations, recipients did not fixate the information about the sources longer when discrepancies were named than without naming of discrepancies. This indicates that participants might not have engaged in sourcing. Previous research, investigating texts from multiple sources frequently showed that conflicts between two different authors led to longer processing times of source information when these sources presented conflicting claims in single texts (Braasch et al., 2012; Rouet et al., 2016) or in multiple text documents (Kammerer et al., 2016) than when they presented consistent claims. This is interpreted as recipients trying to regulate the conflict by engaging in sourcing. My result regarding processing times of source information therefore seems to be inconsistent with results of previous research. However, it can be argued that the source information in my study did not provide relevant information for conflict resolution. In addition, relying solely on eye-tracking to measure sourcing is a limited methodological approach to measure sourcing and additional indicators are helpful (Salmerón et al., 2018).

With regard to conflict regulation strategies, in Study 1, I was further interested whether recipients would engage in information seeking in order to resolve the conflict with additional information. Information seeking was indicated by interest in further information about the painting. I expected that the unresolved conflicts would lead to a perceived knowledge gap and the recipients would be motivated to fill this gap (Loewenstein, 1994; Murayama et al., 2019). Results of Study 1 did not reveal higher interest in further information when discrepancies were named compared to when they were not named. Therefore, my result does not support information seeking as a strategy to resolve the conflict (Graesser et al., 1994). Previous research showed that conflicts can lead to information seeking indicated by asking more information seeking questions after a conflict was encountered (Graesser & McMahen, 1993). However, several explanations for the null

effect of discrepancies on interest are possible. It could be that recipients engaged in another conflict-regulation strategy that was not measured in Study 1. Since the audio-text was commenting on the picture it might have been that the audio-text was perceived to be a more reliable source. Alternatively, I used paintings that are not only perceived as historical documents but also as artworks, theories of art processing could help to explain the missing effect. According to Silvia (2013), interest is triggered by novel information that is easily understood by the individual, while confusion is triggered by novel information that cannot be easily understood. This claim would be in line with our result of interest which, contrary to our expectations, was at least descriptively lowered by naming discrepancies. Based on the different theoretical assumptions, it can be argued that general interest is lowered by the naming of discrepancies, but recipients become more interested to receive specific information helpful to resolve the discrepancy. I investigated the assumption that general interest could be lowered in Study 3 and the assumption that specific interest in information explaining the conflict is enhanced when no other information is provided to explain the conflict in Study 4.

Since the information about the sources in Study 1 were formulated on a general level and were not clearly helpful to reconcile the conflict, I provided information about the sources in Study 2 that either explained or did not explain the conflicts by describing the artists' benevolent intentions. Results of Study 2 revealed that recipients rated trustworthiness of the pictures lower when discrepancies were named and not explained, compared to when discrepancies were not named. As expected, this effect of a lower trustworthiness was compensated by presenting information about the artist's benevolent intentions explaining the discrepancy. This result indicates that recipients engaged in sourcing and used information about the content together with source information presented in the audio-text in order to rate trustworthiness of the painting. This is in line with previous research using multiple text documents showing that recipients consider information about the source, such as benevolence or expertise in order to rate trustworthiness of documents (Gottschling et al., 2019; Kammerer et al., 2016; Thomm & Bromme, 2016).

The results regarding memory did not reveal the expected better memory for pictorial elements when they were presented with discrepancies named compared to without discrepancies named. Literature using single text material suggests that conflicts can lead to an improved memory for the conflicting content (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995). However, in Study 2, an exploratory analysis revealed that memory for the pictorial elements presented always as consistent in all conditions was lower when discrepancies were named than without. Worse memory for coherent content between text and picture due to the presentation of other conflicting information is in line with the findings of Experiment 1 by Schüler (2017). However, in contrast, Experiment 2 by Schüler (2017) and another study using text picture combinations (Schüler, 2019) did not indicate worse memory of consistent content due to the presence of conflicts. It can be argued that conflicts may shift attention to conflicting content and generate a higher cognitive load for the recipient than consistent information (Sweller, 1994). This shift towards conflicting elements including longer processing can sometimes lead to better memory for conflicting content (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995) but not for consistent content (Albrecht & O'Brien, 1993; Hakala & O'Brien, 1995; Schüler, 2017, 2019). However, this might be true when recipients have enough time to process the information. Especially in situations where the consistent information is presented transiently by audio explanations, this shift of attention towards conflicting elements could decrease memory for other information than the conflicting content. This would explain the exploratory result of worse memory when discrepancies were named than when they were not named in Study 2. However, since the result is exploratory the ad hoc explanation remains speculative.

Study 1 and 2 used audio-texts and pictures as documents with the audio-text commenting on the picture as the other document. Due to this commenting role of one document, it might have been that the audio-text was perceived as the more believable document than the picture. It could be argued that therefore recipients did not experience a deep need for conflict regulation in order to decide whom to trust. For this reason, Study 4 was designed to overcome this limitation. In Study 4, the audio-text commented on

consistent and discrepant pictorial elements between two pictures depicting the same event without referring to “reality”. In Study 4, I manipulated the differences in benevolence between the sources of two paintings depicting the same historic event. This paradigm was closely oriented at previous text studies also manipulating source information (Gottschling et al., 2019). Results of Study 4 neither showed a higher number of clicks on the button to read the source information again after the presentation of the paintings nor longer processing of source information when differences between the artists’ benevolence were present than when differences were absent. This indicates that recipients did not engage in more sourcing when information about the source was useful to resolve the conflict than when it was not useful. This result differs from previous research that reported longer processing times of source information when the two sources differed in benevolence than when they did not differ (Gottschling et al., 2019). In Study 4, the differences in processing time of the two paintings, in a task following the presentation, did not differ between the two groups with differences in the artists’ benevolence present and differences in artists’ benevolence absent. Hence, the recipients did not use the information about benevolence in the task where it was helpful to decide on the potentially more trustworthy account. The differences in trustworthiness rating of the two paintings did not differ when differences in artists’ benevolence were present than when differences in artists’ benevolence were absent. Again, this result of Study 4 differs from previous text research. In previous research differences in perceived trustworthiness of documents were higher when the sources were presented with differences in benevolence than without differences in benevolence (Thomm & Bromme, 2016). In Study 4, recipient did not cite the sources more often in a subsequent essay and they did not better remember the source information when differences in artists’ benevolence were presented than when differences in artists’ benevolence were absent. Previous results regarding memory of source information are mixed. One study indicated better source memory when differences in benevolence were present than when they were absent (Thomm & Bromme, 2016). Another study did not report such differences attributing the null effects to ceiling effects due to the high retention rate of the names of the sources indicating

source memory in this study (Gottschling et al., 2019). Because I measured source memory not only by including retention of the artists' names, but also included information relevant for the conflict, I expected to use a more fine-grained method preventing ceiling effects. However, results of Study 4 do not support better source memory when differences in authors benevolence are present than when they are absent. Finally, in Study 4, recipients did not report lower interest in further information about the source or about the conflicting content when differences between artists' benevolence were present than when they were absent.

Based on the previous findings, the null effects of Study 4 regarding differences in source processing, evaluation, source memory, source use and interest between the condition of differences in artists' benevolence being present and differences in artists' benevolence being absent were unexpected. Unfortunately, the manipulation check of Study 4 revealed that differences in trustworthiness of the two artists depicting the same event were small when differences in artists' benevolence were absent as well as when differences in artists' benevolence were present. Although the contrast was significant and the differences between the trustworthiness of the two artists were higher when differences in artists' benevolence were present than when differences in artists' benevolence were absent, the overall ANOVA of differences in the three conditions was not significant. This indicates that the manipulation did not work very well. It could be that a high number of recipients did not believe the manipulation. Considering answers of participants on the task that asked them to decide on the more trustworthy painting in order to describe how the event most likely was in reality, about 17% of the participants in the condition with differences in benevolence did not prefer the documentary painting over the propaganda painting but described both paintings as potentially being biased in their depiction of the event. Alternatively, there might be a lack of expertise in using source information when considering picture documents. Skills in sourcing are mostly trained with text documents (Britt & Aglinskas, 2002) while pictures are rather neglected in history education (Burke, 2001). Therefore, the participants might have ignored source information. In the previous text study,

many participants were able to construct reliable source-content links, resulting in a ceiling effect for memory of the authors name of the respective documents (Gottschling et al., 2019). In my study, only 56% source-content links were correctly remembered, although the name of each artist was stated 5 times during the video together with the respective painting. Additionally, the written text about each artist was given before and after the video. Moreover, only four artists with their respective paintings had to be remembered in my study.

In sum, Study 1 and 2 support generalization of results from text studies to multiple document situations using other very different media formats such as static pictures together with an audio-text directly commenting on the pictures. However, Study 4 indicates that generalization from multiple documents of text to material using other media formats, such as pictures is not trivial and may depend on the design of the learning material as well as on the expertise of the learners with picture documents.

5.2 Results Regarding Pictures as Narratives

In three of four studies I investigated the second aspect of pictures, namely the picture as a narrative telling a story (Study 1, Study 2, Study 3). Until now narrative processing and the resulting transportation are, to the best of my knowledge, not examined with pictures although theories assume that narrative processing is relevant for different kinds of media. Based on the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008), I hypothesized that the naming of discrepancies between the historic event depicted in a picture and the reality as it can be assumed based on todays' historians opinions should violate external realism and thereby lower the recipients experienced transportation. In Study 1 the results revealed lower transportation with discrepancies named, compared to without discrepancies named. However, this was only measured with a single item and I tried to replicate the finding in Study 2 using the six-item short scale (Appel et al., 2015). The results revealed lower transportation with, compared to without naming of discrepancies only for one of three pictures and only if no explanation was given. Since I asked the participants about transportation into the historic event in Study 1 and about

transportation into the depicted event in Study 2, I assessed transportation again in Study 3 using the short scale, but again asking about the real historic event as I did in Study 1. It can be assumed that the real historic event is highly imaginative whereas the depicted event was visible, and the discrepancies should mainly affect the imagination of the real event. However, Study 3 again failed to replicate the results of Study 1. The experienced transportation in Study 3 was equally high when discrepancies were named, compared to when they were not named. In sum, these three studies do only provide weak support for a lower transportation when external realism of a picture is violated by referring to the real historic event.

Previous research reported that attitudes of readers of newspaper article were changed by reading the article, regardless of whether the story was labeled to be true or untrue after reading the article. Hence, the naming of inconsistencies between the facts of the story and the real world did not affect the readers narrative processing. This was also indicated by an absence of an effect of naming inconsistencies on transportation (Green & Donahue, 2011). In contrast to this previous research, the wrong facts in my studies were clearly identifiable by the participants or even presented together with alternative facts about reality (Study 2,3,4). In addition, these facts were presented during the processing of the narrative and not afterwards. However, the naming of discrepancies still did not affect narrative processing indicated by transportation. My result is not in line with the model of narrative processing and engagement, which states that violations of external realism should lower transportation (Busselle & Bilandzic, 2008). Several, explanations are possible. First, I measured transportation offline, after the presentation of the paintings. It could be that an online measurement would reveal an effect on transportation directly after naming discrepancies. Second, it could be that violations of external realism must violate central dimensions of the narrative. Central to the construction of a situational model about the narrative are five dimensions, time, space, protagonist, causality, and intentionality (Zwaan et al., 1995). In my experiments only one central aspect was violated for the narrative of the Washington painting and no central aspect in all other paintings. Discrepancies in my studies

mostly concerned side characters and objects. It might be that violations of external realism need to concentrate on central aspects relevant for the construction of the situational model to lower the experienced transportation. In sum, my results and previous research do not support the claim of the model of narrative processing and engagement (Busselle & Bilandzic) and future research should further examine under which conditions a lowering of transportation due to violations of external realism is valid.

5.3 Results Regarding Pictures as Artworks

I investigated the third aspect, the aesthetic experience and evaluation of the picture in Study 3. I was interested, whether additional information affects subjective understanding, liking, and other aesthetic emotions (surprise, interest, confusion) of viewers of representational art, since effects of additional information are mostly limited to abstract art (Chmiel & Schubert, 2019).

Study 3 revealed that the naming of unexplained discrepancies between the depicted and the real historic event lowered the viewer's liking of the picture compared to not naming discrepancies. In addition, this effect was compensated by information about the artists intentions as an explanation for the discrepancies. Research in the context of aesthetic evaluation of paintings often revealed positive effects of additional information on the liking of artworks (Belke et al., 2010; Bubić et al., 2017; Gerger & Leder, 2015; Millis, 2001; Russell, 2003; Swami, 2013). However, based on theories about aesthetics (Bullot & Reber, 2013; Pelowski et al., 2017; Reber et al., 2004) additional information can affect liking of artworks positively as well as negatively. In Study 3 I observed that additional information affected the liking of representational artworks negatively, if the information confronts the viewer with discrepancies between the representation of the painting and the reality which it seems to represent. This is in line with previous studies showing that semantically mismatching titles can affect the liking of paintings negatively, compared to a control group receiving no titles for representational and abstract paintings (Belke et al., 2010; Gerger & Leder, 2015). Both mismatching titles and the naming of historical inaccuracies confront the

viewer with discrepant information. These discrepancies could decrease the fluency of processing and meaning making and subsequently lower liking (Reber et al., 2004). Discrepancies could also lead to incongruence in the viewer's congruency check as assumed by the VIMAP model (Pelowski et al., 2017). In addition to the lowering of liking due to discrepancies, Study 3 showed that presenting information that explains the conflicts compensated this negative effect of additional information. Only some previous evidence reported small positive effects of additional information in the form of consistent titles on the liking of representational artworks (Belke et al., 2010) while other studies concluded that positive effects might be restricted to abstract art and do not apply to representational art (Leder et al., 2006; Swami, 2013). The result that additional information influences liking indicates that positive effect with additional consistent information may be possible also with representational art and that therefore, the absence of positive effects on liking of representational art in previous studies might actually be explained by ceiling effects and the already high liking of representational artworks when presented without additional information (Chmiel & Schubert, 2019).

In Study 3, neither the naming of discrepancies nor the provision of information to explain them did affect the subjective understanding of the viewer. Previous studies frequently reported positive effects of additional information on subjective understanding of abstract art (Bubić et al., 2017; Leder et al., 2006; Russell, 2003; Swami, 2013). Based on the theories it can be assumed that additional information could also lower the viewer's subjective understanding of the artworks by providing discrepant information that is less easy to comprehend and makes the viewer think about what he or she does not know about an artwork (Bullot & Reber, 2013). Study 3 did not reveal lower subjective understanding with discrepancies named compared to without and subsequently no compensating effect of additional information explaining the discrepancies was possible. Although theories of art evaluation would suggest that subjective understanding and liking of artworks tend to be linked (Bullot & Reber, 2013; Pelowski et al., 2017; Reber et al., 2004), Study 3 did only reveal an effect on liking but not on subjective understanding. This is also in line with

previous research rarely supporting a link between subjective understanding and liking of artworks (Chmiel & Schubert, 2019). Only one study confirmed this link in several experiments (Swami, 2013). Some studies reported effects on subjective understanding but not on liking of artworks (Experiment 2a of Jucker et al., 2014; Leder et al., 2006). One study reported effects on subjective understanding and liking only for semi-abstract paintings. For abstract paintings the subjective understanding was not affected but only liking was affected (Bubić et al., 2017). Many other studies did not simultaneously measure subjective understanding and liking (Chmiel & Schubert, 2019). Unfortunately this includes the two previous studies reporting negative effects of additional information on liking of paintings (Belke et al., 2010; Gerger & Leder, 2015). Therefore, subjective understanding and liking seldom seem to be directly related, and an increased or decreased liking is not necessarily linked to a better or worse subjective understanding.

Lastly with regard to emotions, the results of Study 3 neither support that the naming of discrepancies leads to higher surprise, nor higher confusion and lower interest compared to not naming the discrepancies. Subsequently there was no compensating effect of additional information explaining the discrepancies. Previous research showed that novices tend to be more confused and less interested than experts when viewing complex artworks (Silvia, 2013). But this study did not provide additional information. Some previous studies providing additional information in the form of titles are in line with my findings and also did not reveal positive effects on the emotional experience in general (Bubić et al., 2017) or specific emotions such as interest (Gerger & Leder, 2015; Leder et al., 2006). These studies only used titles and I expected more emotional engagement in my studies by providing longer and more content specific information. However, this was not the case. In our study, liking was highly correlated with interest, surprise, transportation and moderately with subjective understanding. This corresponds with models of art processing claiming a strong link between liking and other aesthetic emotions (Pelowski et al., 2017). However, since the naming of discrepancies and the explanation affected liking but neither subjective understanding nor other aesthetic emotions, how can the effect on liking in Study 3 be

explained? Results of Gerger and Leder (2015) revealed effects of matching and mismatching titles on liking but not on interest. The authors argued that viewers based their liking not on emotional engagement but on fluency. In this sense fluency was higher with matching titles than mismatching titles. This might be the case since laypersons tend to rely their evaluation of artworks on lower stages, while only experts seem to rely on higher order stages of art processing (Mullennix & Robinet, 2018). In line with these authors it can be argued that discrepancies in my study lead to a greater disfluency compared to not naming discrepancies but providing an explanation restored fluency. According to the VIMAP, viewers can rely their evaluation of liking also on bottom-up processing (stage 2 to stage 4), including evaluation of fluency, instead of higher order processes, including meaning making and emotional evaluation. Since laypersons of art were tested in Study 3, the results are in line with fluency theory, but also with the VIMAP and the null effects of naming discrepancies on subjective understanding and other aesthetic emotions than liking can be explained by non-expert viewers relying on lower stages and not engaging in higher stages of art processing.

5.4 Theoretical Implications

The present studies have several implications regarding theoretical models in the field of multiple documents, narrative processing and aesthetic evaluation. In the field of multiple documents, two models were relevant for my research. The framework of multiple document comprehension (Perfetti et al., 1999) describes how readers represent content and source information when encountering multiple text documents. The CSI (Stadtler & Bromme, 2014) focuses on situations where documents make conflicting claims and how readers detect conflicting information and regulate the conflict. Research in the context of these models focused on readers of text documents. Documents in other media formats than written texts where seldom considered. In addition, it was early criticized that various degrees of conflict should be considered in multiple document research. In this sense conflicts can be subtler than applied in most text studies, for example when authors use

different terms describing the same concept (Braasch & Bråten, 2017). Conflicts can also be more explicit when one document directly comments on the other document. In my experiments I used picture documents and accompanying audio-texts explicitly commenting on the pictures. My results support a generalization of effects to multiple documents of mixed media formats and to such an explicit degree of conflicts.

Investigating material not only in the form of written texts, but also in other media formats is necessary since textual documents and combinations of paintings and audio explanations differ in some points relevant for conflict detection and regulation:(1) Text in contrast to pictures is processed linearly from beginning to end and pictures are processed non-linear and selectively (Larkin & Simon, 1987). Non-linear processing might increase the chance that less salient pictorial elements or conflicting pictorial elements can be overlooked more easily (Hegarty et al., 2010) if they are not directly named (Glaser & Schwan, 2015). (2) Texts might be not as immersive as pictures, since pictures represent the facts more vividly. With a combination of text and picture, transportation of the viewers might be increased compared to only texts (Walter et al., 2017). Since more transported viewers tend to believe story claims more and make fewer counterarguments (Green & Brock, 2000), individuals might ignore minor discrepancies of the presented facts more easily (Busselle & Bilandzic, 2008). (3) Written text is less transient than audio-text (Wong et al., 2012), which might increase the chance to detect conflicts in written text. (4) However, written text cannot be presented simultaneously to pictures whereas combinations of picture and audio text can. Simultaneous presentation increases temporal contiguity (Ginns, 2006) and might enhance co-activation of conflicting information and therefore conflict detection compared to a sequential processing of written texts.(5) Lastly, in contrast to textual documents, text and picture combinations or video presentation have the potential for dual coding and therefore might enhance knowledge acquisition (Paivio, 1990).

Despite all these potentially important differences of the research material in the present studies, compared to the previous studies, Study 1 and 2 provide evidence that

generalizations are possible. Since Study 4 could not show generalization effects, it must be noticed that generalizations from text to other media formats are not always trivial. Therefore, researchers should consider that recipients might be less experienced with using sources of pictures than texts due to a lack of training in formal education. In addition, media formats can result in possible boundary condition. My above-mentioned list of potential differences that emerge from media formats can be a fruitful extension of theories in the context of multiple document learning and could be examined systematically by future research. However, my list should does not claim to be exhaustive. A recent study supports that considering boundary conditions of media formats is important. In this study viewers of videos remembered the profession of one source better when it was presented in the video than in written text (Salmerón et al., 2020). This improved memory could be explained by dual coding (Paivio, 1990). Hence, I conclude that the results of my studies support generalization of effects to other media formats. However, based on the above described considerations different media formats might in some combinations limit generalization from text to other media formats and should be systematically investigated by future research.

In Study 2, I tried to disentangle effects of conflicts on memory for congruent information and information either presented as congruent or conflicting. This is necessary since multiple documents often provide some information that is congruent between the documents and some information that is conflicting between the documents. It is reasonable that the presence of conflicts only enhances memory for the conflicting content, due to the fact that previous text research showed that conflicts increase processing times only on the content that is conflicting but not on congruent content (e.g. Albrecht & O'Brien, 1993). Some studies revealed improved memory of conflicting content (Hakala & O'Brien, 1995) but not for the always congruent content (Hakala & O'Brien, 1995; Schüler, 2017, 2019). Results of Study 2 did not reveal better memory for conflicting information. However, an exploratory result indicated that memory for congruent information was worse when conflicts were named than when they were not named. This impairment of memory could have been fostered by a transient presentation of information in my study. It might be that the conflict

drew on limited cognitive capacities (Sweller, 1994) that were missing for processing and memorizing the coherent information of the audio-text. Since conflicting claims, coherent claims, and claims provided by only one side are at the core of multiple document learning, the model should be extended with assumptions on how this can affect memory for information about a complex topic. Although present empirical evidence from single text studies suggest better memory for conflicting information, it would be useful to incorporate other theories to multiple document research when using different media types. For example, based on multimedia theory (Mayer, 2009) it could be assumed that coherent text-picture information is better remembered than conflicting information, since it makes it easier for the recipient to construct a coherent mental model. The exploratory result of Study 2 can provide first evidence that the processing of conflicting text-picture information might need more cognitive capacity than processing only coherent information. Memory might be impaired because more resources are allocated to the processing of conflicting information, which might be lacking for processing coherent information, at least when time is limited, as it is the case with transient audio-texts. This would support the assumption implied by multimedia theory that the construction of a mental model is easier with coherent than conflicting information. Although this argument seems reasonable it relies on an exploratory finding that should be replicated before overestimating it.

The model of narrative comprehension and engagement (Busselle & Bilandzic, 2008) assumes that violations of external realism lower the recipients experienced transportation of a narrative. In my experiments I implemented violations of external realism of narrative paintings by naming discrepancies between the content of the historical painting and the “real” historic event. However, my results of three experiments provide only weak support for the assumption of the model of narrative comprehension and engagement, at least for picture narratives. To the best of my knowledge these were the first studies investigating this assumption of the model by providing clear statements which information of the story is wrong. However, the question remains whether the violations of external realism must violate central dimensions of the narrative. Central to the construction of a situational

model about the narrative are five dimensions, time, space, protagonist, causality, and intentionality (Zwaan et al., 1995). In my experiments the central aspect of time was violated for the narrative of the Washington painting (the crossing of the river Delaware was painted as taking place at daytime while actually it took place at night). All other discrepancies concentrated on the presence or absence of side characters or appearances of persons and objects. It might be that violations of external realism need to be more severely focused on central dimensions of the narrative than it was applied in my studies. In this case, this should be specified by the model of narrative comprehension and engagement.

In the context of aesthetics previous evidence regarding models of art processing and evaluation (Bullot & Reber, 2013; Reber et al., 2004) were mostly limited to abstract art (Chmiel & Schubert, 2019) although the models should apply to abstract as well as representational art. In addition, the models predict positive as well as negative effects of additional information on the liking of art, but evidence is mostly provided for positive effects. Results of Study 3 showed negative effects of additional information on the liking of representational art and that the negative effect can be compensated by further information. Therefore, my studies support these assumptions of the model on positive and negative effects of additional information for representational art. However, the effect of liking was neither paralleled by an effect on subjective understanding nor by an effect on other aesthetic emotions. This, however is surprising since a strong link between liking and subjective understanding or other aesthetic emotions are assumed by the models (Bullot & Reber, 2013; Pelowski et al., 2017). The models should specify if some stages of art processing apply to laypersons of art as well as experts since laypersons might base their liking often on lower stages of art processing such as fluency and not on higher stages associated with emotions and understanding (Gerger & Leder, 2015; Mullennix & Robinet, 2018).

5.5 Practical implications

Pictures of historic events including history paintings are often used in history textbooks. Unfortunately, these pictures are often only used to illustrate the text and are not considered for history learning. In formal educational settings of history learning mostly text documents are used and picture documents tend to be neglected (Burke, 2001). However, recipients need the ability to acquire information from pictures and to consider content of pictures together with source information, otherwise they can be easily misguided. Study 2 revealed that viewers used information about sources of pictures when this information is presented clearly and directly after the conflicting information. In contrast, Study 4 revealed that participants, who encountered conflicting information between two pictures did not use source information presented before and after the audio-visual presentation. Furthermore, in Study 4, participants performed poorly in constructing source-content links indicated by a low memory for the names of the artists of the paintings. Therefore, and due to the widespread use of pictures in communication of information it would be beneficial to consider pictures more often in formal trainings of multiple document strategies to improve learning with multiple documents in other formats than only written text.

Some of the present results can be of interest to museum educators. Museums often provide verbal explanations in the form of personal or audio guides that intend to support the viewers' understanding of the exhibited paintings, the depicted content (in the present studies historic events) and the artist. Therefore, for museums of art and history, it is useful to know that the naming of discrepancies affects the gaze behavior of the viewers. Study 1 revealed that viewers fixate longer on discrepant pictorial elements when they are named as being discrepant to the real historic event, than when they were named as being consistent by a simultaneously presented audio-text. An exploratory result of Study 1 indicated that recipients of a picture and an accompanying audio-text naming the discrepancies often correctly tag the pictorial elements as discrepant and do not mix up correct facts and wrong facts at least shortly after the presentation of the paintings. An exploratory result of Study 2 indicated that the naming of discrepancies can lead to a worse memory for other information, such as the consistent pictorial elements. In addition, if

discrepancies are named it might be beneficial to also provide an explanation for the discrepancies directly after the naming of the discrepancies since this compensated the negative effects of the naming of discrepancies on the evaluation of trustworthiness of the paintings (Study 2). Museum educators and designers of audio guide texts might find these results relevant to their work.

Museums should keep in mind that although research shows frequently positive effects of additional information on liking (e.g. Leder et al., 2006; Swami, 2013), providing additional information is not always beneficial for the aesthetic experience of the viewer but can also lower liking of artworks (Belke et al., 2010; Gerger & Leder, 2015). In Study 3 I showed that this does not only apply to mismatching titles but also to the naming of a painting's historical inaccuracies. The naming of historical inaccuracies is a more valid situation in museums than mismatching titles, since museums often do accompany history paintings with verbal explanations naming these historical inaccuracies. Verbal explanations in contrast to titles can be actively designed by the museum educators. In addition, if discrepancies are named museums should consider providing a subsequent explanation as this can compensate the naming's negative effect on aesthetic liking. However, for practical implication it should be considered that I only investigated laypersons of art and results are likely to be different for art experts. For example, based on the theory (Bullot & Reber, 2013) experts could like the artwork even more when historical inaccuracies are named and not explained, if the expert is able to reach a better artistic understanding on his or her own.

5.6 Strengths, Limitations, and Future Directions

For my research I would first like to highlight the following strengths. The need to consider other media formats than written texts in the context of multiple documents was new and innovative. Pictures are important as documents in history since they can provide historical evidence (Burke, 2001) and also in other fields, such as politics, advertisement of consumer products and natural science. Conflict detection and regulation with pictures and audio-text could differ from written text due to effects of different media formats. Hence, it

was not clear whether effects known from text research would generalize to multiple document situations using such different material. Recent publications of other researchers using video documents (List, 2018; List & Ballenger, 2019; Merkt & Huff, 2020; Salmerón et al., 2020) confirm this need to extend multiple documents research from text to other media formats. My work extends theories and empirical findings in the context of the multiple document models such as the CSI (Stadtler & Bromme, 2014).

In the field of narrative processing and transportation, to the best of my knowledge, my studies were the first that directly manipulated violations of external realism by naming discrepancies between the content of a picture and reality. This provides insights in the context of the model of narrative processing and engagement (Busselle & Bilandzic, 2008). Another strength of my work regards the results about transportation. In the first study, I did observe an effect of discrepancies reducing transportation as it is claimed by the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008). Although the result was in line with the theory, it was limited methodologically because it was only measured with one item. Therefore, I tried to replicate this finding before overestimating limited results. As the result of the first replication was not clear, I again tried to replicate the finding. After three studies I can conclude that support of discrepancies lowering transportation is weak. In empirical research, replications are important especially in times of a replication crisis.

My research is also innovative in the context of aesthetics. Studies mostly used titles or written text information. To the best of my knowledge Study 3 was one of the first using audio-text and supporting positive as well as negative effects of additional information on the aesthetic liking. Audio-texts are often used in museums and in contrast to titles can be actively designed by museum educators. Using such real-world material increases the possibility that effects generalize from the laboratory to the field and enhances ecological validity.

Like all empirical research my work also includes limitations. I used very specific picture material, namely historical paintings. Persons in everyday life mostly see these

paintings in museum or as illustrations in history textbooks and might consider them as art but seldom use them as historical documents. This would be different for other picture material, such as photography, maps, or sketches. Future research could adapt the paradigm I used in Study 4 with photography instead of paintings. This could make the manipulation more believable as recipients might have experience with documentary pictures from newspapers but also fake or propaganda pictures from the Internet. In Study 2 and 3 the material always provided benevolent intentions of the artist to explain the discrepancy. It therefore remains unclear whether the positive effect was caused by the explanation itself or by the information about benevolent intentions. Hence, other types of explanations should be examined with regard to their potential to compensate negative effects of naming discrepancies. The paintings in my studies were presented as pictures on a screen. Therefore, they highly differed in size from the original painting, which could affect transportation.

My research contains also some methodological limitations. First, I tested memory about the content in Study 2 shortly after the presentation of the picture and accompanying audio texts. It would be interesting for future studies to test memory about the content after a longer period of time. Second, in my studies I had to exclude more participants than I expected. For example, in Study 3 nearly 30% of the participants were excluded based on predefined criteria such as prior knowledge of the paintings. This might have been due to using realistic material in my studies. Consequently, for economic reasons, I did not always achieve the sample size I aimed for based on a-priori power calculations. However, the approximated power based on the actual sample sizes was still acceptable (about 60% - 80% across the studies), but readers should keep this in mind when reasoning about the absence of some effects. I also tried to minimize this limitation by the incorporated replications. Lastly, Study 3 and 4 were conducted online. It is more difficult to ensure controlled conditions online than in the laboratory. For example, screen sizes can differ between participants. Also, they might be interrupted during the presentation of the audio-text and did not truly state this at the end of the experiment.

Based on my research, future research could much more consider pictures in the field of scientific learning. For example, in debates about climate change often plots depicting increases in carbon dioxide and world temperature are used for argumentations. In addition, future research could consider effects of different media formats on conflict detection and regulation more systematically. In my work I outlined differences of using picture and audio-text. However, I did not systematically investigate these differences but aimed at collecting evidence for a generalization of results from text research to a situation using media differing in many such aspects from written text. It would also be interesting to examine other explanations or compare uncertain versus certain explanation for discrepancies because such explanations for discrepancies are often speculative and alternative explanations are often discussed in the real world. Lastly, it would be beneficial to use textual narratives about historic events for investigating the assumption of the model of narrative comprehension and engagement that violations of external realism lower the recipient's transportation. The information of pictures could be more vividly imagined than textual information and might not be challenged by textual alternatives.

5.7 Conclusion

The aim of the present studies was to investigate viewers processing, evaluation and memory of pictures accompanied by audio-texts. Thereby the audio provided additional information about the pictures and directly commented on conflicts between the pictures content and the real historic event based on today's historians' opinions or between two pictures. For my investigations I used history paintings and considered relevant theories from three different fields, namely multiple documents, narrative processing and aesthetics. In the context of multiple documents, I conclude that effects known from text material can generalize to material in other media formats such as pictures and audio-text. The explicit naming of discrepancies led to longer processing times on the pictorial content, indicating conflict detection. Viewers engaged in sourcing and used source information together with information about the content to rate trustworthiness of the documents. Considering future

research, researchers should investigate aspects of different media formats as potential boundary conditions for generalization of effects known from text studies as well as trainings to compensate the viewers' limited experience of using pictures as documents. In addition, other pictures than historical paintings should be investigated, such as graphs in science education or photographic pictures.

With regard to narrative processing I conclude that my research only provides weak support for a lowering of transportation when external realism is violated, which is claimed by the model of narrative comprehension and engagement (Busselle & Bilandzic, 2008). Across my studies the power of picture narratives to transport recipients into the narrative was not consistently hampered by pointing out the pictures' wrongly presented facts about the "real" historic event. I would recommend investigating this similarly with text-narratives to draw final conclusions about the model's assumptions.

In the context of aesthetics my results support aesthetic theories that additional information can have positive as well as negative effects on liking of artworks. This provides evidence that additional information can also affect liking of representational art and supports the argument that the present restriction of this effect on abstract art and the absence of effects on representational art could be due to an already high liking of representational art and thus the occurrence of ceiling effects (Chmiel & Schubert, 2019). I did not observe any effects on subjective understanding and other aesthetic emotions and the effects are only valid for laypersons.

In sum, the present studies serve for a better understanding of the processing, evaluation and memory of picture material accompanied by audio-text in the context of multiple documents, narratives and aesthetics. The results are relevant for the theories in the respective fields and provide insights for future research as well as practical implication for education in formal and informal learning settings.

6. References

Albrecht, J. E., & Myers, J. L. (1995). Role of context in accessing distant information during reading. In *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(6), 1459–1468. https://doi.org/10.1037//0278-7393.21.6.1459

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1061–1070. https://doi.org/10.1037//0278-7393.19.5.1061

Appel, M., Gnambs, T., Richter, T., & Green, M. C. (2015). The transportation scale–short form (TS–SF). *Media Psychology*, *18*(2), 243–266. https://doi.org/10.1080/15213269.2014.987400

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior*, *11*(4), 365–374. https://doi.org/https://doi.org/10.1080/10862967909547342

Beker, K., Jolles, D., Lorch, R. F., & van den Broek, P. (2016). Learning from texts: Activation of information from previous texts during reading. *Reading and Writing*, *29*(6), 1161–1178. https://doi.org/10.1007/s11145-016-9630-3

Belke, B., Leder, H., & Carbon, C. C. (2015). When challenging art gets liked: Evidences for a dual preference formation process for fluent and non-fluent portraits. *PLoS ONE*, *10*(8), 1–34. https://doi.org/10.1371/journal.pone.0131796

Belke, B., Leder, H., Strobach, T., & Carbon, C. C. (2010). Cognitive fluency: High-level processing dynamics in art appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(4), 214–222. https://doi.org/10.1037/a0019648

Berlyne, D. E. (1960). Conflict, arousal, and curiosity.

Bilandzic, H., & Busselle, R. W. (2011). Enjoyment of films as a function of narrative

experience, perceived realism and transportability. *Communications*, *36*(1). https://doi.org/10.1515/comm.2011.002

Blanc, N., Kendeou, P., van den Broek, P., & Brouillet, D. (2008). Updating situation models during reading of news reports: Evidence from empirical data and simulations. *Discourse Processes*, *45*(2), 103–121. https://doi.org/10.1080/01638530701792784

Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-Induced Source Comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist*, *52*(3), 167–181. https://doi.org/10.1080/00461520.2017.1323219

Braasch, J. L. G., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, *40*(3), 450–465. https://doi.org/10.3758/s13421-011-0160-6

Bråten, I., & Braasch, J. L. G. (2018). The role of conflict in multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. Mccrudden (Eds.), *Handbook of Multiple Source Use* (pp. 184–201). Routledge. https://doi.org/10.4324/9781315627496

Bråten, I., Salmerón, L., & Strømsø, H. I. (2016). Who said that? Investigating the Plausibility-Induced Source Focusing assumption with Norwegian undergraduate readers. *Contemporary Educational Psychology*, *46*, 253–262. https://doi.org/10.1016/j.cedpsych.2016.07.004

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J. F. (1999). Content integration and source separation in learning from multiple texts. *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso*, 209–233.

Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and

their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instructions, and learning processes* (pp. 276–314). Cambridge University Press. https://doi.org/10.1017/CBO9781139048224

Bubić, A., Sušac, A., & Palmović, M. (2017). Observing individuals viewing art: The effexts of titles on viewers' eye-movement profiles. *Empirical Studies of the Arts*, *35*(2), 194–213. https://doi.org/10.1177/0276237416683499

Bullot, N. J., & Reber, R. (2013). The artful mind meets art history: Toward a psycho-historical framework for the science of art appreciation. *Behavioral and Brain Sciences*, *36*(2), 123–137. https://doi.org/10.1017/S0140525X12000489

Burke, P. (2001). *Eyewitnessing*. Reaktion Books Ltd.

Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, *18*(2), 255–280. https://doi.org/10.1111/j.1468-2885.2008.00322.x

Butler, A. C., Zaromb, F. M., Lyle, K. B., & Roediger III, H. L. (2009). Using popular films to enhance classroom learning: The good, the bad, and the interesting. *Psychological Science*, *20*(9), 1161-1168. https://doi.org/10.1111/j.1467-9280.2009.02410.x

Chmiel, A., & Schubert, E. (2019). Psycho-historical contextualization for music and visual works: A literature review and comparison between artistic mediums. *Frontiers in Psychology*, *10*, 182. https://doi.org/10.3389/fpsyg.2019.00182

Flanagin, A. J., & Metzger, M. J. (2003). The perceived credibility of personal web page information as influenced by the sex of the source. *Computers in Human Behavior*, *19*(6), 683–701. https://doi.org/10.1016/S0747-5632(03)00021-9

Gerger, G., & Leder, H. (2015). Titles change the esthetic appreciations of paintings. *Frontiers in Human Neuroscience*, *9*, 1–10. https://doi.org/10.3389/fnhum.2015.00464

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*.

Yale University Press.

Gilbert, D. T. (1991). How mental systems believe. *American psychologist*, *46*(2), 107–119. https://doi.org/10.1037/0003-066X.46.2.107

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*(6), 511–525. https://doi.org/https://doi.org/10.1016/j.learninstruc.2006.10.001

Glaser, M., Knoos, M., & Schwan, S. (2020). How verbal cues help to see and understand art. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/https://doi.org/10.1037/aca0000372

Glaser, M., & Schwan, S. (2015). Explaining pictures: How verbal cues influence processing of pictorial learning material. *Journal of Educational Psychology*, *107*(4), 1006–1018. https://doi.org/10.1037/edu0000044

Gottschling, S., Kammerer, Y., & Gerjets, P. (2019). Readers' processing and use of source information as a function of its usefulness to explain conflicting scientific claims. *Discourse Processes*, *56*(5–6), 429–446. https://doi.org/10.1080/0163853X.2019.1610305

Graesser, A. C., & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology*, *85*(1), 136–151. https://doi.org/10.1037/0022-0663.85.1.136

Graesser, A. C., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, *95*(3), 524–536. https://doi.org/10.1037/0022-0663.95.3.524

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

https://doi.org/10.1037/0033-295X.101.3.371

Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, *38*(2), 247–266. https://doi.org/10.1207/s15326950dp3802_5

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701–721. https://doi.org/10.1037/0022-3514.79.5.701

Green, M. C., & Brock, T. C. (2002). In the mind's eye: Transportation-imagery model of narrative persuasion. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations.* (pp. 315–341). Lawrence Erlbaum Associates Publishers.

Green, M. C., & Donahue, J. K. (2011). Persistence of belief change in the face of deception: The effect of factual stories revealed to be false. *Media Psychology*, *14*(3), 312–331. https://doi.org/10.1080/15213269.2011.598050

Hakala, C. M., & O'Brien, E. J. (1995). Strategies for resolving coherence breaks in reading. *Discourse Processes*, *20*(2), 167–185. https://doi.org/10.1080/01638539509544936

Hegarty, M., Canham, M. S., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology*, *36*(1), 37–53. https://doi.org/10.1037/a0017683

Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes*, *53*(5–6), 430–456. https://doi.org/10.1080/0163853X.2016.1169968

Jucker, J.-L., Barrett, J. L., & Wlodarski, R. (2014). "I Just Don'T Get it": Perceived artists' intentions affect art evaluations. *Empirical Studies of the Arts*, *32*(2), 149–182.

https://doi.org/10.2190/EM.32.2.c

Kendeou, P., Smith, E. R., & O'Brien, E. J. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology Learning Memory and Cognition*, *39*(3), 854–865. https://doi.org/10.1037/a0029468

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/https://doi.org/10.1037/0033-295X.95.2.163

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100. https://doi.org/https://doi.org/10.1016/S0364-0213(87)80026-5

Lauring, J. O., Pelowski, M., Forster, M., Gondan, M., Ptito, M., & Kupers, R. (2016). Well, if they like it . . . effects of social groups' ratings and price information on the appreciation of art. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(3), 344–359. https://doi.org/10.1037/aca0000063

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489–508. https://doi.org/10.1348/0007126042369811

Leder, H., Carbon, C. C., & Ripsas, A. L. (2006). Entitling art: Influence of title information on understanding and appreciation of paintings. *Acta Psychologica*, *121*(2), 176–198. https://doi.org/10.1016/j.actpsy.2005.08.005

List, A. (2018). Strategies for comprehending and integrating texts and videos. *Learning and Instruction*, *57*, 34–46. https://doi.org/10.1016/j.learninstruc.2018.01.008

List, A., Alexander, P. A., & Stephens, L. A. (2017). Trust but verify: Examining the association between students' sourcing behaviors and ratings of text trustworthiness. *Discourse Processes*, *54*(2), 83–104.

https://doi.org/https://doi.org/10.1080/0163853X.2016.1174654

List, A., & Ballenger, E. E. (2019). Comprehension across mediums: the case of text and video. *Journal of Computing in Higher Education*, *31*(3), 514–535. https://doi.org/https://doi.org/10.1007/s12528-018-09204-9

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75. https://doi.org/https://doi.org/10.1037/0033-2909.116.1.75

Loughlin, S., Grossnickle, E., Dinsmore, D., & Alexander, P. (2015). “Reading” paintings: Evidence for trans-symbolic and symbol-specific comprehension processes. *Cognition and Instruction*, *33*(3), 257–293. https://doi.org/10.1080/07370008.2015.1076822

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children’s vigilance towards deception. *Cognition*, *112*, 367–380. https://doi.org/10.1016/j.cognition.2009.05.012

Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press.

Mayer, R. E., & Fiorella, L. (2014). 12 principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In *The Cambridge handbook of multimedia learning* (Vol. 279). Cambridge University Press.

Merkt, M., & Huff, M. (2020). Does the position of source information for multiple documents matter? Insights from two experiments. *Contemporary Educational Psychology*, *62*, 101900. https://doi.org/https://doi.org/10.1016/j.cedpsych.2020.101900

Millis, K. (2001). Making meaning brings pleasure: The influence of titles on aesthetic experiences. *Emotion*, *1*(3), 320.

Moore, K. M., & West, A. N. (2012). Global perception, meaning, and aesthetic appreciation. *Empirical Studies of the Arts*, *30*(1), 23–38. https://doi.org/10.2190/EM.30.1.d

Mullennix, J. W., & Robinet, J. (2018). Art expertise and the processing of titled abstract art. *Perception*, *47*(4), 359–378. https://doi.org/10.1177/0301006617752314

Murayama, K., FitzGibbon, L., & Sakaki, M. (2019). Process account of curiosity and interest: A reward-learning perspective. *Educational Psychology Review*, *31*(4), 875–895. https://doi.org/10.1007/s10648-019-09499-9

Nell, V. (1988). *Lost in a book: The psychology of reading for pleasure.* Yale University Press.

O'Brien, E. J., & Cook, A. E. (2016). Coherence threshold and the continuity of processing: The RI-Val model of comprehension. *Discourse Processes*, *53*(5–6), 326–338. https://doi.org/10.1080/0163853X.2015.1123341

O'Brien, E. J., & Myers, J. L. (1985). When comprehension difficulty improves memory for text. *Journal of Experimental Psychology Learning Memory and Cognition*, *11*(1), 12–21. https://doi.org/10.1037/0278-7393.11.1.12

O'Brien, E. J., Rizella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1200–1210. https://doi.org/10.1037/0278-7393.24.5.1200

Otero, J., & Graesser, A. C. (2001). PREG: Elements of a model of question asking. *Cognition and Instruction*, *19*(2), 143–175. https://doi.org/10.1207/S1532690XCI1902_01

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press.

Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017). Move me, astonish me… delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, *21*, 80–125.

https://doi.org/10.1016/J.PLREV.2017.02.003

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Towards a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99–122). Lawrence Erlbaum Associates, Inc.

Pozzer-Ardenghi, L., & Roth, W. (2005). Making sense of photographs. *Science Education*, *89*(2), 219–241. https://doi.org/https://doi.org/10.1002/sce.20045

Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, *35*(8), 2019–2032. https://doi.org/https://doi.org/10.3758/BF03192934

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*(4), 364–382. https://doi.org/10.1207/s15327957pspr0804_3

Rouet, J.-F., Le Bigot, L., de Pereyra, G., & Britt, M. A. (2016). Whose story is this? Discrepancy triggers readers' attention to source information in short narratives. *Reading and Writing*, *29*(8), 1549–1570. https://doi.org/10.1007/s11145-016-9625-0

Russell, P. A. (2003). Effort after meaning and the hedonic value of paintings. *British Journal of Psychology*, *94*(1), 99–110. https://doi.org/10.1348/000712603762842138

Salmeron, L., Gil, L., & Bråten, I. (2018). Using eye-tracking to assess sourcing during multiple document reading: A critical analysis. *Frontline Learning Research*, *6*(3), 105–122. https://doi.org/10.14786/flr.v6i3.368

Salmerón, L., Sampietro, A., & Delgado, P. (2020). Using internet videos to learn about controversies: Evaluation and integration of multiple and multimodal documents by primary school students. *Computers & Education*, *148*, 103796. https://doi.org/10.1016/j.compedu.2019.103796

Saux, G., Britt, A., Le Bigot, L., Vibert, N., Burin, D., & Rouet, J. F. (2017). Conflicting but

close: Readers' integration of information sources as a function of their disagreement. *Memory and Cognition*, *45*(1), 151–167. https://doi.org/10.3758/s13421-016-0644-5

Saux, G., Ros, C., Britt, M. A., Stadtler, M., Burin, D. I., & Rouet, J.-F. (2018). Readers' selective recall of source features as a function of claim discrepancy and task demands. *Discourse Processes*, *55*(5–6), 525–544. https://doi.org/10.1080/0163853X.2018.1463722

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., & Scherer, K. R. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLOS ONE*, *12*(6), e0178899. https://doi.org/10.1371/journal.pone.0178899

Schüler, A. (2017). Investigating gaze behavior during processing of inconsistent text-picture information: Evidence for text-picture integration. *Learning and Instruction*, *49*, 218–231. https://doi.org/10.1016/J.LEARNINSTRUC.2017.03.001

Schüler, A. (2019). The integration of information in a digital, multi-modal learning environment. *Learning and Instruction*, *59*, 76–87. https://doi.org/10.1016/J.LEARNINSTRUC.2017.12.005

Segal, E. M. (1995). Narrative comprehension and the role of deictic shift theory. In F. Duchan, G. A. Bruder, & L. E. Hewitt (Eds.), *Deixis in narrative: A cognitive science perspective* (pp. 3–17). Hillsdale, NJ: Lawrence Erlbaum Associates.

Silvia, P. J. (2009). Looking past pleasure: Anger, confusion, disgust, pride, surprise, and other unusual aesthetic emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(1), 48–51. https://doi.org/10.1037/a0014632

Silvia, P. J. (2013). Interested experts, confused novices: Art expertise and the knowledge emotions. *Empirical Studies of the Arts*, *31*(1), 107–115. https://doi.org/10.2190/EM.31.1.f

Smith, M., Breakstone, J., Wineburg, S., & Wineburg, S. (2019). History assessments of thinking : A validity study. *Cognition and Instruction*, *37*(1), 118–144. https://doi.org/10.1080/07370008.2018.1499646

Solso, R. L. (2003). The psychology of art and the evolution of the conscious brain. In *ACS Applied Materials and Interfaces* (Vol. 7, Issue 45). MIT Press. https://doi.org/10.1021/acsami.5b07841

Specker, E., Forster, M., Brinkmann, H., Boddy, J., Pelowski, M., Rosenberg, R., & Leder, H. (2018). The Vienna Art Interest and Art Knowledge Questionnaire (VAIAK): A unified and validated measure of art interest and art knowledge. *Psychology of Aesthetics, Creativity, and the Arts. Advance Online Publication.* https://doi.org/10.1037/aca0000205

Stadtler, M., & Bromme, R. (2014). The content–source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 379–402). MIT Press.

Stadtler, M., Scharrer, L., & Bromme, R. (2012). Does relevance matter in comprehending scientific conflicts from multiple documents? Evidence from online and offline-data. In E. de Vries & K. Scheiter (Eds.), *Staging knowledge and experience: How to take advantage of representational technologies in education and training? Proceedings of the EARLI SIG 2 Meeting* (pp. 202–204).

Stadtler, M., Scharrer, L., Skodzik, T., & Bromme, R. (2014). Comprehending multiple documents on scientific controversies: Effects of reading goals and signaling rhetorical relationships. *Discourse Processes*, *51*(1–2), 93–116. https://doi.org/10.1080/0163853X.2013.855535

Stein, N. L., & Trabasso, T. (1985). The search after meaning: Comprehension and

comprehension monitoring. In F. J. Morrison, C. Lord, & D. Keating (Eds.), *Applied developmental psychology* (pp. 33–58). Academic Press.

Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*, *20*(3), 192–204. https://doi.org/10.1016/j.learninstruc.2009.02.001

Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction*, *31*(2), 176–203. https://doi.org/10.1080/07370008.2013.769994

Swami, V. (2013). Context matters: Investigating the impact of contextual information on aesthetic appreciation of paintings by Max Ernst and Pablo Picasso. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(3), 285–295. https://doi.org/10.1037/a0030965

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7*(14), 4. https://doi.org/10.1167/7.14.4

Thomm, E., & Bromme, R. (2016). How source information shapes lay interpretations of science conflicts: Interplay between sourcing, conflict explanation, source evaluation, and claim evaluation. *Reading and Writing*, *29*(8), 1629–1652. https://doi.org/10.1007/s11145-016-9638-8

van Boxtel, C., & van Drie, J. (2012). “That’s in the time of the romans!” Knowledge and strategies students use to contextualize historical images and documents. *Cognition and Instruction*, *30*(2), 113–145. https://doi.org/10.1080/07370008.2012.661813

Vorderer, P. (1992). *Fernsehen als Handlung: Fernsehfilmrezeption aus motivationspsychologischer Perspektive [Watching television as action: Reception of TV*

movies from the perspective of motivational psychology]. Edition Sigma.

Walter, N., Murphy, S. T., Frank, L. B., & Baezconde-Garbanati, L. (2017). Each medium tells a different story: The effect of message channel on narrative persuasion. *Communication Research Reports*, *34*(2), 161–170. https://doi.org/10.1080/08824096.2017.1286471

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*(1), 73–87. https://doi.org/http://dx.doi.org/10.1037/0022-0663.83.1.73

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*(6), 449–457. https://doi.org/https://doi.org/10.1016/j.learninstruc.2012.05.004

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. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

7. Electronic Supplementary Material

Appendix A: Washington Crossing the Delaware: Picture and Label in Study 1

Schematic presentation of the material depicting the painting “Washington Crossing the Delaware” with its respective label (only used in Experiment 1) in the lower left corner.

Label

The label states in German: “Washington Crossing the Delaware” (1851) is an oil painting by the American artist Emanuel Leutze. He painted it in Germany about 70 years after the event.”

Source of the Leutze painting:

https://upload.wikimedia.org/wikipedia/commons/9/95/Washington_Crossing_the_Delaware_by_Emanuel_Leutze%2C_MMA-NYC%2C_1851.jpg?uselang=de

Appendix B: Washington Crossing the Delaware – Audio Text in Study 1

The complete audio-text accompanying the painting “Washington Crossing the Delaware” in the first experiment. The original text was in German. The differing sentences of the conditions are marked with A and B.

(Segment 1: General information) The painting depicts the leader of the American Revolutionary Army, General George Washington, on December 26, 1776 while crossing the Delaware River on the East Coast of America. The crossing was part of the surprise attack against the British and gave the Americans a key victory in the War of Independence.

(Segment 2: Main character) In the left middle of the painting, George Washington is standing. He is wearing a saber and a tricorn hat. After the American War of Independence, Washington became the first president of the United States of America.

(Segment 3: Non-discrepant pictorial element) Behind Washington, a man is standing wearing a blue uniform and a tricorn hat. It is James Monroe who left university to join the war for Americas independence. Later, he was elected to become the fifth president of the United States.

(Segment 4: Non-discrepant pictorial element) Far right in the painting, a man is kneeling, wearing a green jacket and a fur cap and holding the rudder. It is a local hunter who knew the area around the Delaware River. The local hunters helped with their knowledge of the place and in addition were good marksmen.

(Segment 5: Discrepant pictorial element) In the middle of the painting, the American flag with the white and red stripes and the stars on a blue ground is waving in the wind. At that time, many different flags were used in the Revolutionary Army.

A: The flag depicted was one of the best known. *(Without discrepancy)*

B: The flag depicted however was not yet invented. *(With discrepancy)*

(Segment 6: Discrepant pictorial element) In the upper left corner the sun is already high up in the sky. The crossing of the Delaware was taking many hours because of the bad weather.

A: However, the crossing was carried out and finished during daytime. *(Without discrepancy)*

B: However, the crossing was carried out and finished during the night. *(With discrepancy)*

Appendix C: Washington Crossing the Delaware – Audio Text in Study 2

An example of the audio-text for one pictorial element of the Leutze painting in the second experiment (original text was in German). For the four conditions, we manipulated whether the discrepancy was named by the text or not and then whether an explanation was given in the subsequent text or not:

At the left top of the painting, the sun can be seen. The position of the sun shows that it is late in the morning.

Followed by either A or B:

A: Without discrepancy: The crossing was actually completed by day in the light of the sun.

B: With discrepancy: But the crossing was actually done at night without the light of the sun.

Followed by either C or D:

C: Without explanation: Leutze painted the scene with a shining sun and depicted the colors magnificently. Looking back, he saw the crossing as a triumphal event with historical significance and thought that he could never make this great significance clear to the viewer with any form of representation.

D: With explanation: Leutze painted the scene with a shining sun so that he could depict the colors more magnificently. Looking back, he saw the crossing as a triumphal event with historical significance and thought that he could only make this great significance clear to the viewer with this form of representation.

Appendix D: An Example of an Answer Given for the Painting Washington Crossing the Delaware in the Retention Test in Study 1.

Bitte nenne alle Bildelemente des Gemäldes (z.B. Landschaftselemente, Gegenstände und einzelne Personen), die Du noch erinnern kannst. Beschreibe sie möglichst eindeutig.

2. Pferde auf einem hinteren Boot (sind dunkel, eins hell, eine in rot gekleidete Frau in der Mitte des Bootes von Washington, links auf Washingtons Boot ein Mann der mit seinem Fuß versucht sich von den Eisblöcken im Wasser abzustoßen, die amerikanische ... weiße Sterne auf dunkelblauem Hintergrund + rot-weiß Flagge in der Mitte Washingtons Boot (leicht verdeckt mit Holzstab), ein in grün gekleideter Jäger, der das Boot steuert, Washington links neben der Flagge, Flagge wird gehalten von einem Mann, der später 5. Präsident Amerikas geworden ist; zwei Boote im Hintergrund, grau bewölkter Himmel, Sonne kommt links oben leicht durch

Bitte ergänze deine Beschreibung durch eine grobe Zeichnung. Du kannst diese auch beschriften, damit klar ist um wen oder was es sich handelt.

Appendix E: Example Text for information about the artist Edward Penny in Study 4.

After the general information either a, b, or c followed depending on whether the artists intention was named as propagandistic, documentary, or not named. The intention is underlined. The information asked in the memory task is in bold.

Edward Penny (1714-1791) was a British painter. He studied art first with a private teacher in London and later in Rome. In his early years he mostly painted small-piece portraits later he dedicated himself to elaborate history paintings. He was a founding member of the Royal Academy of Arts in London. His paintings can be seen today in the Ashmoleon Museum in Oxford, among others. His painting “The Death of General Wolfe” is one of his most famous works. Edward Penny painted it in 1763 **during the Seven Years War**. In this war Great Britain fought against France among others.

1. Penny's painting depicting Wolfe and his companions was intended to <u>inspire</u> the **British audience** with a <u>propagandistic</u> depiction.
2. Penny's painting depicting Wolfe and his companions was intended to <u>inform</u> the **British audience** with a <u>documentary</u> representation
3. Penny's painting depicting Wolfe and his companions was exhibited in England for a **British audience**.